UNIVERSITY OF MICHIGAN
OCT 12 1954
SOCIAL SCIENCE LIBRARY

Psychometrika

A JOURNAL DEVOTED TO THE DEVELOPMENT OF PSYCHOLOGY AS A QUANTITATIVE RATIONAL SCIENCE

THE PSYCHOMETRIC SOCIETY - ORGANIZED IN 1935

VOLUME 19
NUMBER 3
SEPTEMBER
1 9 5 4

PSYCHOMETRIKA, the official journal of the Psychometric Society, is devoted to the development of psychology as a quantitative rational science. Issued four times a year, on March 15, June 15, September 15, and December 15.

SEPTEMBER, 1954, VOLUME 19, NUMBER 3

Published by the Psychometric Society at 1407 Sherwood Avenue, Richmond 5, Virginia. Entered as second class matter at the Post Office of Richmond, Virginia. Editorial Office, Department of Psychology, The University of North Carolina, Chapel Hill, North Carolina.

Subscription Price: The regular subscription rate is $14.00 per volume. The subscriber receives each issue as it comes out, and, upon request, a second complete set for binding at the end of the year. All annual subscriptions start with the March issue and cover the calendar year. All back issues are available. Back issues are $14.00 per volume (one set only) or $3.50 per issue, with a 20 per cent discount to Psychometric Society Members. Members of the Psychometric Society pay annual dues of $7.00, of which $6.30 is in payment of a subscription to *Psychometrika*. Student members of the Psychometric Society pay annual dues of $4.00, of which $3.60 is in payment for the journal.

Application for membership and student membership in the Psychometric Society, together with a check for dues for the calendar year in which application is made, should be sent to

MAHLON V. TAYLOR, JR.
American Institute for Research, 410 Amberson Avenue, Pittsburgh 32, Pa.

Payments: All bills and orders are payable in advance.

Checks covering membership dues should be made payable to the *Psychometric Society*.

Checks covering regular subscription to *Psychometrika* (for non-members of the Psychometric Society) and back issue orders should be made payable to the *Psychometric Corporation*. The checks, notices of change of address, and business communications should be sent to

JOHN W. FRENCH, Treasurer, Psychometric Society and Psychometric Corporation
Educational Testing Service
P.O. Box 592
Princeton, New Jersey

Articles on the following subjects are published in *Psychometrika:*

(1) the development of quantitative rationale for the solution of psychological problems;
(2) general theoretical articles on quantitative methodology in the social and biological sciences;
(3) new mathematical and statistical techniques for the evaluation of psychological data;
(4) aids in the application of statistical techniques, such as nomographs, tables, work-sheet layouts, forms, and apparatus;
(5) critiques or reviews of significant studies involving the use of quantitative techniques.

The emphasis is to be placed on articles of type (1), in so far as articles of this type are available.

(Continued on the back inside cover page)

Psychometrika

CONTENTS

AN ANALYTICAL METHOD FOR SIMPLE STRUCTURE . . . 173
L. L. Thurstone

A METHOD FOR THE STUDY OF INTERSTIMULUS SIMILARITY 183
Clyde H. Coombs

TWO KINDS OF FACTOR ANALYSIS? 195
Raymond E. Hartley

A SIMPLE PROOF OF A PERSONNEL CLASSIFICATION THEOREM 205
Hubert E. Brogden

VERIFICATION OF HYPOTHESIZED FACTORS IN ONE HUNDRED AND FIFTEEN OBJECTIVE PERSONALITY TEST DESIGNS 209
R. B. Cattell, S. S. Dubin and D. R. Saunders

NOTE ON CONSTRUCTION OF AN ITEM ANALYSIS TABLE FOR THE HIGH-LOW-27-PER-CENT GROUP METHOD 231
C. T. Fan

CHANGES IN FACTOR STRUCTURE OF A COMPLEX PSYCHOMOTOR TEST AS A FUNCTION OF PRACTICE 239
Edwin A. Fleishman and Walter E. Hempel, Jr.

OSGOOD AND SUCI'S MEASURE OF PATTERN SIMILARITY AND Q-TECHNIQUE FACTOR ANALYSIS 253
James M. Sakoda

A COMPUTING CHART FOR THE POINT BISERIAL CORRELATION COEFFICIENT 257
Harvey F. Dingman

A NOTE ON ITEM ANALYSIS WITH AN ELECTRONIC COMPUTER 261
James Ziegler

VOLUME NINETEEN SEPTEMBER 1954 NUMBER 3

PSYCHOMETRIKA—VOL. 19, NO. 3
SEPTEMBER, 1954

AN ANALYTICAL METHOD FOR SIMPLE STRUCTURE

L. L. THURSTONE

UNIVERSITY OF NORTH CAROLINA

One of the most difficult parts of multiple-factor analysis for the student to learn is the rotation of the coordinate axes to simple structure. The complete graphical methods have been the most dependable. The work of interpreting these diagrams is not only tedious but it requires the development of a skill that many otherwise competent students are unable to master. The writer has tried many times to develop an analytical method which would eliminate the graphical methods of successive approximation. The method to be described here seems at last to be successful and practically feasible. It is equally applicable to unimodal and bimodal hyperplanes so that no restrictive assumptions need to be made about the positive or negative signs of the factorial components.

The methods of multiple-factor analysis were developed originally for the problem of isolating the primary mental abilities, but these methods have been found useful also in other fields. One of the most difficult parts of multiple-factor analysis for the student to learn is the rotation of the coordinate axes to simple structure. The complete graphical methods have been the most dependable. These methods require a set of $\frac{1}{2}r(r-1)$ plots to be made for each successive adjustment of the coordinate planes where r is the number of factors. Each plot represents graphically the relation between a pair of columns of the factor matrix F. The work of interpreting these diagrams is not only tedious but it requires the development of a skill that many otherwise competent students are unable to master. As a result there are many published multiple factor studies in which the authors attempt to interpret a rotated factor matrix without making plots of all pairs of columns.

The writer has tried many times to develop an analytical method which would eliminate the graphical methods of successive approximation. The method to be described here seems at last to be successful and practically feasible. It is equally applicable to unimodal and bimodal hyperplanes so that no restrictive assumptions need to be made about the positive or negative signs of the factorial components. If this new method continues to be as successful as it has been in a number of current applications, it should make obsolete the graphical successive approximation procedures on which we have been dependent for some years.

The new method is a single-plane method. By this we mean that each

coordinate hyperplane is determined separately. In these methods it has usually been necessary to start with a hypothesis that a specified sub-group of test vectors lie in a hyperplane of dimensionality $(r - 1)$ where r is the rank of the corresponding correlation matrix. The procedure was then to make successive guesses about which test vectors lie in the same hyperplane. The desired reference vector would then be the normal which defines the hyperplane. This element of uncertainty about the sub-groups can evidently be eliminated.

The method will be described for the three-dimensional case, but it is applicable to a factor problem of any rank. In developing this method we have explicitly avoided any form of solution which involves a characteristic equation. Such a solution usually requires the numerical evaluation of latent roots which is a tedious process because the order is often as large as ten or twelve or even fifteen.

The spherical model is represented in Figure 1 in which the small rings

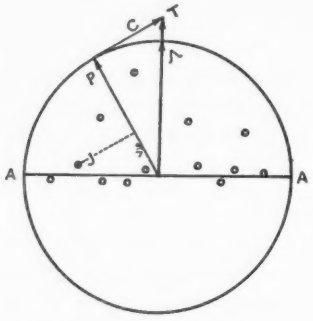


FIGURE 1

denote the termini of extended test vectors on the surface of the unit sphere. Inspection of the spherical model shows nine points $(A—A)$ that over-determine a plane. Four test vectors in Figure 1 evidently do not lie in the same plane. In this case the sub-group of nine test vectors can easily be selected by inspection but in the n-dimensional case it is necessary to use other means for identifying the plane because physical models cannot be made in more than three dimensions.

The present method starts with an arbitrarily selected unit trial vector P with direction cosines ρ_m . This may be a unit vector in one of the co-ordinate axes of an orthogonal factor matrix or it may be preferably a normalized test vector. The problem is to adjust this trial vector P to some nearby position Λ which is the normal that defines the plane A.

A criterion function ϕ_p is defined as

$$\phi_p = \sum_j w_{jp} v_{jp}^2 , \tag{1}$$

where v_{jp} is the scalar product JP and where J is a test vector. Since P is a unit vector, v_{jp} is also the projection of J on P as shown in Figure 1. The criterion function gives a weight w_{jp} to each test vector J. It is desired to find the mean principal axis for the whole set of weighted test vectors.

The weight w_{jp} for each test J is determined solely by the projection of J on P, or

$$w_{jp} = f(v_{jp}). \tag{2}$$

It is desired to give maximum weights to those test vectors that are nearly orthogonal to P and to give minimum weights to those test vectors that have large projections on P. If there is an overdetermined hyperplane whose normal is in the vicinity of P, the maximum projection v_{jp} of a test vector on P in that plane is not likely to exceed, say, .60. A zero weight may be given to each test vector whose projection v_{jp} exceeds .60. Maximum weights are given to those test vectors whose projections v_{jp} are near zero. A simple set of weights that has been found effective is given in Table 1. The weights

TABLE 1

Weights for Tests

v_{jp}	w_{jp}
.00 – .09	6
.10 – .19	5
.20 – .29	4
.30 – .39	3
.40 – .49	2
.50 – .59	1
.60 – 1.00	0

w_{jp} are always positive. The value of w_{jp} is independent of the sign of v_{jp}. The question of uniqueness of the solution will be discussed in a later paragraph.

The trial vector P is to be adjusted with a correction vector C to the new vector T where

$$P + C = T \tag{3}$$

with the condition that C shall be orthogonal to P. Then

$$PC = 0. \tag{4}$$

The direction numbers of T may be denoted t_m . The criterion function of (1) can be expressed for the new vector T in the form

$$\phi_t = \sum_j w_{jp} v_{jt}^2 , \tag{5}$$

where v_{jt} is the scalar product

$$v_{jt} = JT. \tag{6}$$

The numerical values of v_{jt} will be larger than the projections on T because T exceeds unit length. The direction numbers of the test vector J are a_{j1} , a_{j2} , a_{j3} , and these are given in the rows of the factor matrix F. The scalar product JT can then be written in the form

$$v_{jt} = a_{j1}t_1 + a_{j2}t_2 + a_{j3}t_3 . \tag{7}$$

Expanding (5) by (7), we have

$$\phi_t = \sum_j w_{jp}(a_{j1}t_1 + a_{j2}t_2 + a_{j3}t_3)^2. \tag{8}$$

The partial derivatives are then

$$\begin{aligned}
\frac{\partial \phi_t}{\partial t_1} &= 2t_1 \sum_j w_{jp}a_{j1}^2 + 2t_2 \sum_j w_{jp}a_{j1}a_{j2} + 2t_3 \sum_j w_{jp}a_{j1}a_{j3} \\
\frac{\partial \phi_t}{\partial t_2} &= 2t_1 \sum_j w_{jp}a_{j1}a_{j2} + 2t_2 \sum_j w_{jp}a_{j2}^2 + 2t_3 \sum w_{jp}a_{j2}a_{j3} \\
\frac{\partial \phi_t}{\partial t_3} &= 2t_1 \sum_j w_{jp}a_{j1}a_{j3} + 2t_2 \sum_j w_{jp}a_{j2}a_{j3} + 2t_3 \sum w_{jp}a_{j3}^2 .
\end{aligned} \tag{9}$$

The conditional equation could be written as in (4) but it will be computationally simpler to impose the same geometrical condition by the restriction that the scalar product PT shall be unity.* Then

$$\psi = PT - 1 = 0 \tag{10}$$

or

$$\psi = \rho_1 t_1 + \rho_2 t_2 + \rho_3 t_3 - 1 = 0, \tag{11}$$

where ρ_i are the direction numbers of P. The partial derivatives are then

$$\frac{\partial \psi}{\partial t_1} = \rho_1 ; \qquad \frac{\partial \psi}{\partial t_2} = \rho_2 ; \qquad \frac{\partial \psi}{\partial t_3} = \rho_3 . \tag{12}$$

It is desired to minimize the criterion ϕ_t . The normal equations are in the form

$$\frac{\partial \phi}{\partial t_m} - \beta \frac{\partial \psi}{\partial t_m} = 0, \tag{13}$$

*The use of this alternative conditional equation was suggested to the writer by Dr. Ledyard Tucker of Princeton.

where the subscript m refers to the orthogonal axes of F, and β is a Lagrange multiplier. Substituting (9) and (12) in (13) and assembling the terms in rectangular notation, we have

$$\begin{array}{|c|c|c|} \hline \sum_i w_{ip}a_{i1}^2 & \sum_i w_{ip}a_{i1}a_{i2} & \sum_i w_{ip}a_{i1}a_{i3} \\ \hline \sum_i w_{ip}a_{i1}a_{i2} & \sum_i w_{ip}a_{i2}^2 & \sum_i w_{ip}a_{i2}a_{i3} \\ \hline \sum_i w_{ip}a_{i1}a_{i3} & \sum_i w_{ip}a_{i2}a_{i3} & \sum_i w_{ip}a_{i3}^2 \\ \hline \end{array} \cdot \begin{array}{|c|} \hline t_1 \\ \hline t_2 \\ \hline t_3 \\ \hline \end{array} = \begin{array}{|c|} \hline \tfrac{1}{2}\beta\rho_1 \\ \hline \tfrac{1}{2}\beta\rho_2 \\ \hline \tfrac{1}{2}\beta\rho_3 \\ \hline \end{array} \qquad (14)$$

A $\quad$ t $\quad$ $(\frac{1}{2}\beta)\rho$

This equation can be written in the matrix form

$$At = (\tfrac{1}{2}\beta)\rho, \qquad (15)$$

where t and ρ are column vectors.

The values of t_m can then be written explicitly in the relation

$$t = (\tfrac{1}{2}\beta)A^{-1}\rho. \qquad (16)$$

First we compute the column vector u from

$$u_m = A^{-1}\rho_m . \qquad (17)$$

It is not necessary to compute the Lagrange multiplier β because U, T, and Λ are colinear. To find the direction cosines of the desired unit vector Λ we need only normalize U. Then

$$\Lambda = kU, \qquad (18)$$

where

$$k = \frac{1}{\sqrt{\sum u_m^2}}, \qquad (19)$$

so that

$$\Lambda = \frac{1}{\sqrt{\sum u_m^2}}\, U. \qquad (20)$$

If it is desired to find T, we can write

$$T = SU \qquad (21)$$

and, recalling that $PT = 1$ from (10), we have

$$S(PU) = 1, \qquad (22)$$

so that

$$S = \frac{1}{(PU)'} \tag{23}$$

and hence

$$T = \frac{1}{(PU)} U. \tag{24}$$

If the Lagrange multiplier is desired, we write from (16) and (17)

$$T = (\tfrac{1}{2}\beta)U, \tag{25}$$

and by (24) and (25) we have

$$(\tfrac{1}{2}\beta)U = \frac{1}{(PU)} U, \tag{26}$$

so that

$$\beta = \frac{2}{(PU)'}. \tag{27}$$

From these relations it can also be shown that

$$\beta = \frac{2}{\sqrt{\sum u^2 (P\Lambda)}}. \tag{28}$$

The reader may wonder whether a unique solution for the desired normal to an overdetermined hyperplane can be found with a rather arbitrary set of weights w_{ip}. If an overdetermined hyperplane exists in the test configuration the plane will be identified even with considerable liberties in choosing the arbitrary function $w_{ip} = f(v_{ip})$. If an overdetermined hyperplane does not exist, the solution will not be unique. The question whether the plane exists in the configuration will be determined in each case by the number of vanishingly small projections v_{ip} and preferably also by a final set of plots, namely, one for each pair of columns of the rotated factor matrix V. This final set of plots should be included in publications so that the reader can judge the degree of overdetermination of each hyperplane.

Computational procedure

Each element of the matrix A is a product of three numbers such as $w_{ip}a_{i1}a_{i2}$. The elements a_{i1} and a_{i2} are obtained from the orthogonal factor matrix F. A second factor matrix F_w can be prepared with the elements $w_{ip}\,a_{im}$. Then we can express the matrix A in (14) as the matrix product

$$F_w'F = A. \tag{29}$$

This requires the preparation of a new matrix F_w for each hyperplane that is to be determined.

An alternative procedure is to compute a matrix B with elements $a_{jm}a_{jn}$ where the subscripts m and n represent a pair of columns of F. There are $\frac{1}{2}r(r-1)$ such pairs and hence as many columns of B. This matrix is computed only once for a multiple-factor problem. For each hyperplane that is to be determined one must multiply each row by the weight w_{jp}. It is not yet certain which computational procedure is the more advantageous.

After the matrix has been determined its inverse A^{-1} is computed. An inverse of order r must be computed for each hyperplane. The next step is to compute u_m by (17) and Λ by (20), which gives the desired reference vector. The projections $v_{j\lambda}$ are then computed. If a large number of vanishingly small projections $v_{j\lambda}$ are found, the hyperplane is probably overdetermined. The final step is to plot a graph for each pair of columns of the oblique factor matrix V. Inspection of these diagrams shows whether the hyperplanes are acceptable for interpretation.

An important advantage of this method over current methods of rotation is that one can assert with confidence whether a hyperplane is overdetermined. One might find that a certain number of hyperplanes are ready for interpretation and that other specified hyperplanes are indeterminate. With present methods the investigator can only say that he has not found a particular overdetermined hyperplane. In such a situation a more experienced investigator might find it. By the new method here described this type of uncertainty seems to be removed.

One could start with a trial unit vector P in each of the r orthogonal coordinate axes of the factor matrix F. The first centroid axis can be expected to be an unsatisfactory starting point for P because it may require several adjustments. A better procedure is probably to start with a normalized test vector as a unit trial vector P. This should preferably be a test which has some relatively high correlations in the correlation matrix R and which also has an appreciable number of relatively low coefficients. By the present method we find one of the reference vectors Λ_1. One of the test vectors which has a low value for $v_{j\lambda}$ on Λ_1 is normalized to become the trial vector P_2. It is adjusted by the present method to become the second reference vector Λ_2. Then we select a test vector which has low projections on both Λ_1 and Λ_2 to be normalized into the third trial vector P_3. A projection is judged to be low when its absolute value is low relative to the rest of the test projections on the same reference vector. This procedure is continued until the number of reference vectors so determined is equal to the number of factors in the original factor matrix F. The test projections on each reference axis Λ_p constitute a column of the oblique factor matrix V. All pairs of columns of V are then plotted in $\frac{1}{2}r(r-1)$ diagrams and these are examined to ascertain whether they show an overdetermined structure.

Occasionally a trial vector P will be adjusted into a reference vector A that has already been found, but this situation should be rare with the procedure for selecting the successive trial vectors as here described.

The weighting schedule described in Table 1 should fit most of the factorial studies with psychological tests. A more generally applicable plan for setting the weights w_{jp} is to examine the distribution of values of v_{jp} to determine its range of absolute values. The total range of absolute values can be divided into a number of equal intervals, or roughly equal. The number of intervals may be from 4 to 6 or 7. The lowest absolute values of v_{jp} are assigned the maximum weight. The highest absolute values of v_{jp} are given zero weight. In this way a hyperplane that is clearly defined by the test vectors and which is roughly orthogonal to the trial vector P will be identified.

The method is new and it needs experimental try-out on factor matrices of different kinds and sizes to determine empirically whether it is subject to some limitations that are not yet known. The author will appreciate correspondence with those who may care to try it on current problems.

Illustrative Problem

The present solution for simple structure has been tried on several problems with good results. One of these was a four-dimensional problem of anthropometric measurements (**1**). Each of the hyperplanes was located in one trial. The average discrepancy between the previous simple structure solution and the present solution was about .01.

A five-dimensional example will be used here. It was a fictitious problem in which the correlations were produced from a given factor matrix (**2**, p. 230). The correlation matrix was factored and the resulting factor matrix F (**2**, p. 231) was used as the starting point for the present method of locating the hyperplanes. This orthogonal factor matrix F_0 is reproduced here as Table 3.

The first step in the computing is to select one of the test vectors for the first trial. Test 1 was selected. It can be expected that the best results will be obtained by selecting a test which has an appreciable number of low correlations and also at least a few relatively high correlations. In Table 6 this test is identified as $J = 1$ and its five direction numbers are recorded in the first column. These are also the five values of a_{jm} in the first row of F_0 in Table 3. Normalizing the vector $J = 1$ we get the unit vector P_1 which is, of course, colinear with J.

The next step is to obtain the projections v_{jp} of the test vectors on P_1. These are obtained from the rows of F_0 and the column vector P_1 and they are recorded in Table 2. The next column shows the weight w_{jp} that is assigned to each test by Table 1.

TABLE 2

Test	v_{jp}	w_{jp}
1	.855	0
2	.316	3
3	.175	4
4	.245	4
5	.762	0
6	.140	5
7	.000	6
8	.175	4
9	-.001	6
10	.468	1
11	.000	6
12	.106	5
13	.843	0
14	.280	3
15	.281	3
16	.000	6
17	.469	1
18	.656	0
19	.000	6
20	.561	0
21	.000	6
22	.655	0
23	.643	0
24	.387	2
25	-.001	6
Σ	8.015	

TABLE 3 Matrix F_0

Test	I	II	III	IV	V
1	510	-188	529	325	-224
2	368	489	536	-076	378
3	652	474	-026	-304	-084
4	510	306	370	-475	153
5	596	-273	373	320	023
6	611	180	-209	166	452
7	358	-345	-300	102	542
8	517	-030	035	048	685
9	536	074	-557	261	223
10	689	-059	-263	454	-132
11	670	-083	-464	-312	047
12	570	-070	002	-569	132
13	436	-394	394	394	-394
14	506	656	254	056	213
15	637	-436	-111	-304	038
16	536	-347	-280	-422	238
17	555	-521	-044	309	256
18	518	-366	270	-026	-419
19	358	-118	-074	-666	-226
20	559	070	-070	447	-447
21	358	443	-443	246	-246
22	652	-129	057	180	-514
23	670	330	240	330	-162
24	581	-072	246	-476	-203
25	492	399	-471	-004	-330
Σ	13.445	-010	-006	004	-001

TABLE 4 Matrix F_w

Test	I	II	III	IV	V	Σ
1	0	0	0	0	0	0
2	1.104	1.467	1.608	-0.228	1.134	5.085
3	2.608	1.896	-0.104	-1.216	-0.336	2.848
4	2.040	1.224	1.480	-1.900	0.612	3.456
5	0	0	0	0	0	0
6	3.055	0.900	-1.045	0.830	2.260	6.000
7	2.148	-2.070	-1.800	0.612	3.252	2.142
8	2.068	-0.120	0.140	0.192	2.740	5.020
9	3.216	0.444	-3.342	1.566	1.338	3.222
10	0.689	-0.059	-0.263	0.454	-0.132	0.689
11	4.020	-0.498	-2.784	-1.872	0.282	-0.852
12	2.850	-0.350	0.010	-2.845	0.660	0.325
13	0	0	0	0	0	0
14	1.518	1.968	0.762	0.168	0.639	5.055
15	1.911	-1.308	-0.333	-0.912	0.114	-0.528
16	3.216	-2.082	-1.680	-2.532	1.428	-1.650
17	0.555	-0.521	-0.044	0.309	0.256	0.555
18	0	0	0	0	0	0
19	2.148	-0.708	-0.444	-3.996	-1.356	-4.356
20	0	0	0	0	0	0
21	2.148	2.658	-2.658	1.476	-1.476	2.148
22	0	0	0	0	0	0
23	0	0	0	0	0	0
24	1.162	-0.144	0.492	-0.952	-0.406	0.152
25	2.952	2.394	-2.826	-0.024	-1.980	0.516

TABLE 5

$v_{j\lambda}$	(P_λ)
.813	.8
.047	.0
.000	.0
.009	.0
.697	.7
-.011	.0
-.044	.0
-.012	.0
-.032	.0
.478	.5
-.060	.0
-.039	.0
.897	.9
.043	.0
.238	.3
-.067	.0
.460	.5
.675	.7
-.045	.0
.599	.6
.003	.0
.682	.7
.525	.5
.276	.3
-.015	.0

TABLE 6

	J = 1	P_1	U	A
I	.510	.596	.5010	.4548
II	-.188	-.220	-.4118	-.3738
III	.529	.619	.5415	.4915
IV	.325	.380	.5190	.4711
V	-.224	-.262	-.4801	-.4358
Σ	.952	1.113	.6696	.6078

TABLE 7 Matrix A

	1	2	3	4	5
1	21.0509	2.3610	-6.7304	-5.8926	4.7429
2	2.3610	8.0543	.5800	1.5167	-1.6917
3	-6.7304	.5800	8.7632	-1.0145	-.0237
4	-5.8926	1.5167	-1.0145	9.2449	.6082
5	4.7429	-1.6917	-.0237	.6082	7.5789
Σ	15.5318	10.8203	1.5746	4.4627	11.2146

TABLE 8 Matrix A^{-1}

	1	2	3	4	5
1	.2650	-.1883	.2433	.2414	-.2265
2	-.1883	.2705	-.1848	-.1974	.1935
3	.2433	-.1848	.3401	.2366	-.2115
4	.2414	-.1974	.2366	.3349	-.2213
5	-.2265	.1935	-.2115	-.2213	.3340
Σ	.3349	-.1065	.4237	.3942	-.1318

The entries of each row j of F_0 are then multiplied by the corresponding weight w_{jp} and the products are recorded in the matrix F_w in Table 4.

The entries in the symmetric matrix A of Table 7 are then obtained from Tables 3 and 4. The matrix A can be expressed as the matrix product $A = F_0'F_w$. The computational work is done on pairs of columns of F_0 and F_w . The inverse A^{-1} is then computed and recorded as in Table 8.

The matrix product $U = A^{-1}\rho$ is then computed from Table 8 and the column vector P_1 in Table 6. The resulting column vector U is recorded in Table 6. Normalizing U we have the direction cosines of the unit reference vector Λ. This is the adjusted position of the initial trial reference vector P.

The scalar products $v_{j\lambda}$ are then computed from Table 3 and the column vector Λ in Table 6. The resulting values of $v_{j\lambda}$ are recorded as shown in Table 5. In the present case we know the exact values of $v_{j\lambda}$, which are shown as column (P_5) in Table 5. It can be seen that the agreement between the values obtained by the present method of rotation and the correct values is remarkably close and that only one trial was sufficient. So far we have not found it necessary to use a second trial because of slow convergence. If that should be necessary we would use Λ as the new unit trial vector P_2 and proceed exactly as in the first trial to find an adjusted vector Λ_2 and so on.

Professor Thomas Jeffrey has prepared the numerical example that has been described here, and he expects to design IBM machine methods by which the new analytical method can be used advantageously with minimum time and labor.

REFERENCES

1. Thurstone, L. L. Factor analysis and body types. *Psychometrika*, 1946, **11**, 15–22.
2. Thurstone, L. L. Multiple-factor analysis. Chicago: Univ. Chicago Press, 1947.

Manuscript received 10/14/53

PSYCHOMETRIKA—VOL. 19, NO. 3
SEPTEMBER, 1954

A METHOD FOR THE STUDY OF INTERSTIMULUS SIMILARITY*

CLYDE H. COOMBS

UNIVERSITY OF MICHIGAN

A generalized method for collecting data on interstimulus similarity is presented and its special cases evaluated by use of information theory. A method of analyzing the data by the Unfolding Technique is presented which permits the study of the latent attribute structure underlying the similarity of stimuli for a single individual.

I. *Introduction*

For certain purposes it may be desirable to determine the relations among a set of stimuli with respect to their similarity. In certain instances, where the similarity and differences are ascribable to a single common latent attribute, the stimuli may be scaled in any of a number of different ways (**1**, **5**, **6**). The data utilized to construct such a scale, here called a stimulus scale, are either single judgments from each of a number of different judges or many experimentally independent replications of judgments from a single judge. The former are usually used when the stimuli are identifiable and there might thereby be experimental dependence in the replication of a single individual. The latter are necessary when the scale to be constructed is entirely a personal one and not common to other individuals. The Law of Comparative Judgment provides the best method now available for constructing a stimulus scale at the level of interval scale measurement but requires that the successive judgments be experimentally independent and that the stimuli be indiscriminable to some extent.

In certain instances the problem of constructing a stimulus scale for a single individual, when the stimuli are identifiable and 100 per cent discriminable, arises. This problem occurs in constructing scales of the utilities of objects and scales of psychological probability. In each such instance the stimuli may be clearly identifiable, and perfectly discriminable, and the scales may be unique to an individual. Constructing such scales is the problem for which the methods presented in this paper have been developed.

The methods developed here are an adaptation of the Unfolding Technique as it is applied to J scales (**3**, **4**). In the case of a J scale, the different individuals are conceived of as standing on the J scale (the judges' respective "ideals") and evaluating each pair of stimuli as to which is "nearer" (pre-

*This study was carried out under Office of Naval Research Contract Nonr 374 (00) NR 041-011.

ferred). The judge's order of the stimuli from most preferred to least preferred is called an I scale and represents the J scale folded at the judge's ideal. The data, consisting of a set of I scales, may be unfolded to recover the J scale at the level of an ordered metric which may be interpreted as a common latent attribute underlying the preferences.

The problem in a stimulus scale is that the judges are not distributed on the continuum with the stimuli. In the case of a utility scale, for example, the locus of the stimuli on the scale corresponds to their "preferability" or their "desirability" to the individual.

To obtain metric relations between stimuli on the stimulus scale the method of collecting data may be modified to permit the use of the Unfolding Technique. What is necessary is a method of collecting data which will make the individual "stand" on the stimulus scale at a number of different places and evaluate the stimuli in order of increasing distance from each. For each position an I scale is obtained and the set of these may be unfolded to recover the stimulus scale with metric relations.

The general method of collecting data is called the Method of Cartwheels, primarily because it provides a ready terminology for the elements. All of the methods which are special cases of this general method differ only in the *amount* of information they provide; the *kind* of information is the same for all. The stimuli are always presented in sets of three or more in any of several arrangements and the basic judgment operation which the subject performs is to rank order the *distances* between pairs of stimuli. As far as I know, the first person to recognize and make use of such data was Richardson (**8**), followed by Klingberg (**7**) and more recently by Torgerson (**10**), all of whom applied the Law of Comparative Judgment to the analysis of such data. The method of analysis described here is quite different.

In the next section a special case of the Method of Cartwheels is presented, called the Method of Similarities, and the method of analysis is described in detail. The generalized method is then presented and some of its special cases pointed out. The analysis of data collected by any of the special cases is clear from the treatment of the Method of Similarities, as the various special cases differ only in the amount of information they provide.

II. *The Method of Similarities*

The stimuli are presented to the individual in triads, and he is asked to judge which two are most alike and which two are least alike in each triad. Then each triad is decomposed into three paired comparisons, each paired comparison becoming an element used in constructing a different I scale. The method is illustrated with a hypothetical example of five stimuli, $ABCDE$, presented in ten triads as shown in Figure 1.

The position of the stimuli in the triad and the serial order of the triads are randomized in an experiment. An individual is instructed to respond

to each of the ten triads indicating in each which pair of stimuli is most alike and which pair is least alike. This is the most general form of the instructions; variations are described below.

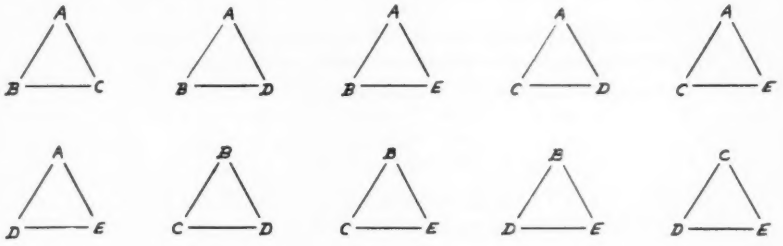
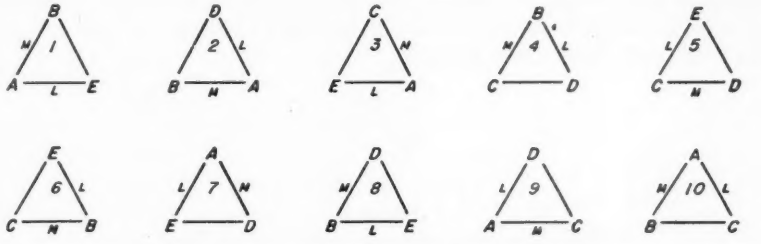


FIGURE 1
The Ten Triads of Five Stimuli

To illustrate the method of analyzing the data, assume the judgments obtained to be those presented in Table 1, and let the hypothesis be that the five stimuli lie on a common latent attribute. It is desired to test this hypothesis and, if sustained, to recover the latent attribute with the stimuli

TABLE 1
Sequence of Triads and the Response of the Judge*



*M signifies the judgment of "most alike," and L signifies the judgment of "least alike."

ordered on it and with order relations on the distances between pairs of stimuli.

Analysis of the data begins by decomposing each triad into three paired comparisons and tabulating them as in Table 2. For example, the first triad is judged as shown in Figure 2. The decomposition of this triad into three paired comparisons proceeds as follows. Imagine that the individual were standing on the stimulus scale where stimulus A is and that he had been asked whether stimulus B or E were nearer A. Clearly, he has answered this

hypothetical question, "B is nearer than E," i.e., $B \cdot > E$. This paired comparison is recorded in column A, the stimulus from whose position the individual was judging.

TABLE 2

Tabulation of Paired Comparisons from Decomposition of Triads*

A	B	C	D	E
1. B ▸ E	1. A ▸ E	3. A ▸ E	2. B ▸ A	1. B ▸ A
2. B ▸ D	2. A ▸ D	4. B ▸ D	4. C ▸ B	3. C ▸ A
3. C ▸ E	4. C ▸ D	5. D ▸ E	5. C ▸ E	5. D ▸ C
7. D ▸ E	6. C ▸ E	6. B ▸ E	7. A ▸ E	6. C ▸ B
9. C ▸ D	8. D ▸ E	9. A ▸ D	8. B ▸ E	7. D ▸ A
10. B ▸ C	10. A ▸ C	10. B ▸ A	9. C ▸ A	8. D ▸ B
I Scales ABCDE	BACDE	CBADE	DCBAE	EDCBA

* ▸ signifies the implied judgment, "stimulus __ is nearer to the stimulus at the head of the column than stimulus __ "

This procedure is continued, imagining that the individual is standing on each stimulus in turn. Thus, standing on stimulus B, he has said that A is nearer than E, $A \cdot > E$, and this is tabulated in column B. Finally, standing on stimulus E, he has said that B is nearer than A, $B \cdot > A$, which is tabulated in column E. Each triad is so decomposed and each paired comparison tabulated in its respective column. The data in Table 1 have been tabulated in Table 2, and each paired comparison has been identified by number with the triad from which it was obtained.

FIGURE 2

Hypothetical Response Pattern for the First Triad

The paired comparisons in a column of Table 2 are experimentally independent and, if they are transitive, they may be reduced to a rank order of the stimuli in order of increasing distance from the stimulus at the head of the column. In these hypothetical data, this is the case and the paired comparisons have been reduced to the rank orders given at the bottom of their respective columns in Table 2.

The problem then is to see if these rank order scales, which are formally equivalent to I scales, satisfy the necessary conditions for unfolding into a common latent attribute (**3**, **4**). These data meet the necessary conditions. For example, two of the I scales, and only two, are mirror images of each

other and thus yield the rank order of the stimuli on the latent attribute, *ABCDE*. They contain no metric information because they do not fold the scale. The other three *I* scales do contain metric information and are listed in Table 3 with their respective metric analyses. Consider the *I* scale *BACDE*.

TABLE 3

Metric Implications from Within I Scales*

BACDE	⇒	$\overline{BC} > \overline{AB}$
CBADE	⇒	$\overline{CD} > \overline{AC}$
DCBAE	⇒	$\overline{DE} > \overline{AD}$

* The bar over a pair of stimuli signifies the distance between them.

This represents the latent attribute folded at stimulus *B*. Hence, the fact that *A* comes before *C* implies that the distance $\overline{BC} > \overline{AB}$. In a similar manner the other two *I* scales have been analyzed and the metric implications are given in Table 3.

The metric information from the different *I* scales is not necessarily consistent or transitive. They must be examined for consistency and transitivity before they can be accepted as sustaining the hypothesis of a single latent attribute with these metric relations. The metric relations are summarized in Figure 3, and they are seen to be consistent and transitive.

FIGURE 3

Composite Partial Order of the Metric Information in the Data

Thus, from these data one would have obtained a stimulus scale with the stimuli in the order *ABCDE* and with the metric relations given in Figure 3. This is as far as the data takes one. The problem of interpreting (identifying) the attribute, and of saying which end is which is an exercise for the experimenter's judgment. In the case of utility and of psychological probability, this is usually self-evident or may be obtained readily from the subject.

There are variations in the instructions that may be used with this method, appropriate for different purposes. It may be desirable to make the latent attribute explicit and ask which are most alike in brightness, in attitude toward the church, or whatever is relevant.

When the latent attribute is explicit, then it is only this subspace of the total stimulus space that is being investigated. One may study an individual's conception of such attributes as leadership, authority, or aesthetic merit, by the structural relations of the stimuli in such a subspace. When the latent attribute is left implicit, the relative similarity of stimuli and the attributes underlying similarity may be investigated, providing a new approach to the study of equivalence of stimuli.

In scaling utilities, we have used instructions like which two are hardest to choose between, and which two are easiest. In scaling psychological probabilities, we have instructed the subject to indicate which is most likely and which is least likely, and to place the intermediate statement nearer one of the two extremes.

Information theory is of assistance in evaluating various methods of collecting data. For example, in the Method of Similarities, each triad may be responded to in six different ways. The three "distances" in each triad are rank ordered by the two judgments "most" and "least" and, as the three distances may be ranked six different ways, there are six different response patterns. Assuming these to be equally likely, the number of bits of information contained in a single triad is $\log_2 6 = 2.58$. For n stimuli the number of triads is $\binom{n}{3}$ so the total number of bits is $2.58\binom{n}{3}$. In decomposing a triad into three paired comparisons, this is making three bits of information out of 2.58, and this increment represents the amount of information that has been added by imposing transitivity on paired comparison judgments within a triad. While there is experimental dependence among these three paired comparisons, the data are analyzed by associating paired comparisons from different triads which are experimentally independent.

A weakness of the method lies in the lack of replication and when there is error or vacillation in judgment, the I scales may not satisfy the necessary conditions for unfolding into a common latent attribute. The method is highly vulnerable and really constitutes a stringent criterion of whether a common latent attribute exists for an individual over these stimuli. As most human behavior seems to involve some "random" elements it was found desirable to generalize this method to provide replication and hence an

evaluation of inconsistency of judgments. The generalization of the Method of Similarities we call the Method of Cartwheels.

III. *The Method of Cartwheels*

The name Method of Cartwheels covers a great variety of methods which might be illustrated as follows. Of the n stimuli, p are presented at a time with not more than one in the center, or hub, and the others around the rim, as in Figure 4. The instructions are similar to those for the Method of

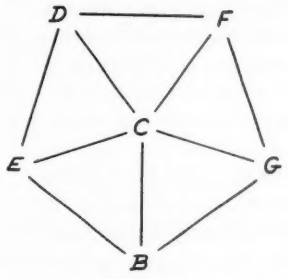


FIGURE 4
Illustrating the Method of Cartwheels

Similarities, except that the subject is asked to indicate which pair is most alike, which pair is next most alike, and so on, to which pair is least alike. A line in the figure, as a spoke, or on the rim, designates a pair of stimuli to be included in the ranking.

Clearly, with n stimuli, the number presented at a time, p, must satisfy $3 \leqq p \leqq n$. With $p = 3$, there are two variations. (See Figure 5.)

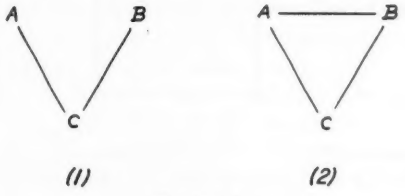


FIGURE 5
(1) The Method of Propellers
(2) The Method of Similarities

In (1), Figure 5, there is a stimulus at the hub and two spokes. The individual judges whether A or B is more like C. This is a direct paired comparison of the two distances. Every set of three stimuli is presented with

each stimulus at the hub in turn; this is spoken of as permuting hubs. Clearly, there is a full degree of freedom for each paired comparison in this method. The paired comparison from each such presentation is tabulated in the column corresponding to the hub, as was done in Table 2 with the Method of Similarities, where each stimulus in turn is regarded as the hub. Because of the number and variety of methods, we have taken to giving names to some of the more frequently used methods. This one is called the Method of Propellers.

In (2), Figure 5, the three stimuli are all on the rim and there is no hub. This is the Method of Similarities already discussed in detail. This method is slightly less powerful than the Method of Propellers when evaluated as follows. The 2.58 bits of information in a presentation in the Method of Similarities are distributed over the three paired comparisons obtained from the decomposition of the triad. Hence, on the average, each paired comparison conveys $(1/3)(2.58) = .86$ bits of information, and there is no replication of the paired comparisons in different triads. In the Method of Propellers each paired comparison conveys 1.00 bit of information. The latter method, however, takes longer to administer, as each set of three stimuli must be presented three times, permuting the hub, whereas in the Method of Similarities a set of three is presented only once. This is another illustration of the general principle that the value of data is measured by the effort required to collect it.

For all $p \geqq 4$ there are further variations:

(1) The stimuli may be all on the rim with no hub (rims only); e.g., one presentation of four stimuli may be arranged as shown in Figure 6. Each set of four may be presented in three different arrangements with each

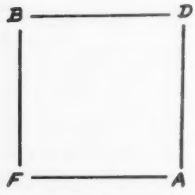


FIGURE 6
Another Variation of the Method of Cartwheels

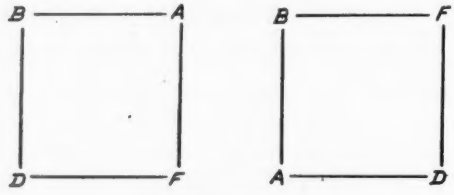


FIGURE 7
Other Arrangements of the Same Four Stimuli

stimulus diagonally opposite each of the others in turn. For example, the four stimuli above may also be presented as shown in Figure 7.

The instructions again are to rank order the distances indicated, from smallest (most alike) to greatest (least alike). The four rank ordered distances may be decomposed into six paired comparisons, four of which are tabulated in the appropriate columns as in Table 2. The other two paired comparisons

are paired comparisons of disjoint distances and are tabulated separately. For example, in a presentation let the rank ordering from most alike to least alike be as illustrated by Figure 8. Imagine again that the individual stands

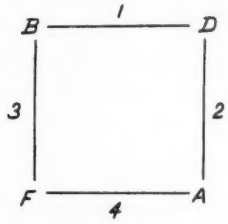


FIGURE 8
A Hypothetical Response Pattern

on each stimulus in turn and judges the relative similarity of the two adjacent stimuli. Then, in the manner of Table 2, these judgments would be tabulated as shown in Table 4. In addition the metric relations $\overline{AD} < \overline{BF}, \overline{AF} > \overline{BD}$

TABLE 4
Tabulation of Paired Comparisons
From Decomposition of the Response Pattern Given in Figure 8

A	B	...	D	...	F
D ▷ F	D ▸ F	...	B ▸ A	...	B ▸ A
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.

are also implied in the rank ordering. The number of presentations of stimuli to which an individual responds in this method is $3\binom{n}{4}$, each set of four stimuli being presented in three arrangements. Each paired comparison, tabulated as in Table 2, is replicated $(n - 3)$ times.

A partial evaluation of this method may be made by applying information theory, as was done to the Method of Similarities. Here the amount of information in a presentation is $\log_2 24 = 4.57$. The judgments in a presentation are decomposed into six paired comparisons, so the 4.57 bits of information in a presentation may be regarded as being distributed over the six paired comparisons. On the average, then, each paired comparison conveys .76 bits of information. As this is cumulative over successive presentations if they are experimentally independent, the number of units of information gathered on any given paired comparison, essentially a measure of the amount of replication, is $.76(n - 3)$. This measure may be interpreted as an indication of the effectiveness of the method in controlling error or inconsistency.

(2) One of the stimuli may be at the hub and all others on the rim, and this has two variations: (a) spokes only, permuting hubs; and (b) both rims and spokes, permuting hubs; e.g., as shown in Figure 9, the instructions

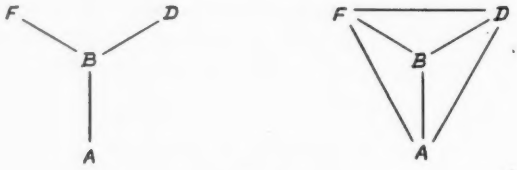


FIGURE 9
Two More Variations for Presenting Four Stimuli at a Time

always being to rank order the pairs or distances indicated from most alike to least alike.

Evaluation of these methods by information theory may proceed as follows. For spokes only, permuting hubs, each presentation yields $\log_2 6 = 2.58$ bits of information, assuming each alternative response pattern is equally likely. These 2.58 units are distributed over 3 paired comparisons, so each receives on the average .86 bits from a single presentation. A given paired comparison is replicated $n - 3$ times in this method, so each paired comparison accumulates a total of .86 $(n - 3)$ bits of information. From the point of view of controlling inconsistency, this method appears more powerful than the preceding.

For the method $p = 4$, both spokes and rims, permuting hubs, the number of alternative response patterns to one presentation is $6! = 720$. So the amount of information in a response pattern is $\log_2 720 = 9.97$. Distributing these over the 15 paired comparisons which a rank ordering of 6 elements decomposes into, each paired comparison may be valued at .66 bits. But in this method each paired comparison is replicated 2 $(n - 3)$ times, so each paired comparison accumulates 1.32 $(n - 3)$ bits of information. This method is clearly the most powerful in this respect of any of those presented here. In fact, we have found that this method taxes the subject too greatly, and increasingly so with increasing complexity of stimuli.

These examples of $p = 3$ and $p = 4$ illustrate some of the variations possible in presenting p stimuli at a time. The major variations may be classified as spokes only, rims only, or both, as illustrated above. The variations differ in their theoretical effectiveness for controlling inconsistency of judgment. With sufficient replications on each paired comparison, inconsistency of paired comparison judgments may be controlled and the dominant preference in each pair determined. The analysis of these paired comparisons into rank ordered I scales and their unfolding proceeds as with

the Method of Similarities if the necessary conditions are satisfied. The result, then, is a scale with the stimuli ordered and with some of the distances between them ordered—an ordered metric.

IV. *General Comments*

The context in which these methods have been presented is the construction of a unidimensional scale for the stimuli. The data may sometimes require a multidimensional structure. Criteria for the number of dimensions necessary to account for the data have been developed by Bennett (**2**), but procedures for recovering a J space are as yet incompletely developed.

It is interesting to note that data collected by these methods may be analyzed also by Torgerson's multidimensional psychophysical scaling method involving an extension of the Law of Comparative Judgment (**10**).

It would be interesting to collect data from a single judge and analyze them by both methods. Application of the Law of Comparative Judgment requires many replications of each paired comparison to permit a reasonably reliable normal curve transformation on the percentages. One of the more powerful versions of the Method of Cartwheels would yield a sufficient number of replications itself, or a less powerful method may be used with replication of the method. We have found the latter to be more feasible because the more powerful versions of the Method of Cartwheels are beyond the judges' capacity to execute and hence generate noise. The scale obtained from Torgerson's method would be an interval scale, the scale from the method of analysis described here would be an ordered metric; but they could be compared.

The methods presented here are most appropriate for the development of stimulus scales which may be peculiar to individuals, and for testing whether different individuals have the same latent structure for a given set of stimuli. They may also prove useful in the study of equivalence of stimuli, generalization, and transfer. By virtue of providing an independent source of metric information, these methods will also be useful in the study of various attributes of response as possible psychological distance functions, e.g., amplitude, latency, and consistency.

REFERENCES

1. Attneave, Fred. A method of graded dichotomies for the scaling of judgments. *Psychological Review*, 1949, **56**, 334–340.
2. Bennett, J. F. A method for determining the dimensionality of a set of rank-orders. Unpublished doctor's dissertation, University of Michigan, 1951.
3. Coombs, C. H. Psychological scaling without a unit of measurement. *Psychological Review*, 1950, **57**, 145–158.
4. Coombs, C. H. A theory of psychological scaling. *Engineering Research Bulletin*, (University of Michigan Press), 1952, No. 34.

5. Edwards, A. L. and Thurstone, L. L. An internal consistency check for scale values determined by the method of successive intervals. *Psychometrika*, 1952, **17**, 169–180.
6. Guilford, J. P. Psychometric methods. New York: McGraw-Hill, 1936.
7. Klingberg, F. L. Studies in measurement of the relations among sovereign states. *Psychometrika*, 1941, **6**, 335–352.
8. Richardson, M. W. Multidimensional psychophysics. *Psychological Bulletin*, 1938, **35**, 659–660.
9. Thurstone, L. L. A law of comparative judgment. *Psychological Review*, 1927, **34**, 273–286.
10. Torgerson, W. S. Multidimensional scaling: I. Theory and method. *Psychometrika*, 1952, **17**, 401–419.

Manuscript received 8/3/53
Revised manuscript received 10/26/53

PSYCHOMETRIKA—VOL. 19, NO. 3
SEPTEMBER, 1954

TWO KINDS OF FACTOR ANALYSIS?*

RAYMOND E. HARTLEY

UNIVERSITY OF MIAMI

In descriptive factor analysis, the factor matrix is merely an alternate mode of expression of the meaning expressed by the correlation matrix. In inferential factor analysis, the factor matrix expresses material inferred from the correlation matrix but not contained in the correlation matrix. A major problem has been whether or not the latter type is possible. Granit's color cone study is interpreted as a factor study in the hope of showing that inferential factor analysis actually exists. It is argued that proof of these and other points in factor theory can be accomplished only by logical analysis.

It would probably be acknowledged by most that there is more than one kind of factor analysis. Indeed, it sometimes seems that there are almost as many variations as there are students of factor analysis. On the other hand most factor analysts may be monistic in the sense that they would classify all these variations as fairly minor modifications on the same theme.

In contrast to this position it is the purpose of this paper to re-examine the dualistic position that there are at least two major varieties of factor analysis. This suspicion that factor analysts tend to be of two types is not new to this paper. Burt (1), for example, discusses factor analysis as "systematic description" and as "causal explanation." He is, however, monistic in the sense that he regards description as the basic function of factor analysis and causal explanation as impossible. In like fashion Kelley remarks that:

> There is no search for timeless, spaceless, populationless truth in factor analysis; rather, it represents a simple, straightforward problem of description in several dimensions of a definite group functioning in definite manners, and he who assumes to read more remote verities into the factorial outcome is certainly doomed to disappointment (**5**, p. 120).

This seems to indicate both a belief that a descriptive function is basic for factor analysis and that another, illegitimate variety of factor analysis has also been proposed. It is this second, questionable variety of factor analysis that is the chief interest of this paper. It will be held here that the dismissal of this second variety, herein called "inferential factor analysis," as illegitimate has been too hasty. The force of the present argument will rest entirely upon the construction of a semi-fictitious factor study which will, apparently,

*This article was read before the Florida Psychological Association, 1953. It is based upon a portion of a dissertation submitted by the writer to the department of psychology of the University of Chicago in partial fulfillment of the requirements for the Ph.D. degree.

be at once inferential but not metaphysical. I realize that such an argument cannot demonstrate the legitimacy of inferential factor analysis, but I hope that it can lend it plausibility.

Descriptive Factor Analysis

In order to sharpen the distinction between inferential and descriptive factor analysis let us first characterize "descriptive" factor analysis. The designation "descriptive" is used in order to be in harmony with what seems to be usual terminology. It is supposed, furthermore, that no questions exist concerning the legitimacy of this form of factor analysis.

The use of an analogy may aid the explanation. It is the draftsman's job, when given an object or piece of machinery, to make an exact representation of the original object. His task is descriptive in that his drawing records the essential features of the original so that the original can be recreated from the drawing by a machinist. In like fashion, the factor matrix retains all the essential features of the correlation matrix. The correlation matrix can be recreated (computed) from the factor matrix by anyone who knows the rules.

The techniques of drawing the draftsman uses are conventional in the sense that they have become standardized or customary by virtue of long usage or formal agreement among engineers and draftsmen. In different situations, furthermore, the draftsman has different techniques. In one drawing he might use the familiar three-view drawings. In another he might utilize cutaway sectional views. In a case where he prefers greater pictorial quality he might use isometric projection and in another, where the mechanism is complicated and the audience inexperienced, he might use the so-called "exploded" drawing.

Correspondingly, the descriptive factorist has available a variety of possible kinds of solutions and devices that he may decide upon. In order to emphasize one aspect of the correlation matrix he may decide upon one factor solution; to emphasize another aspect he might choose a different solution. As with the draftsman's techniques the various factorial solutions are arrived at because they are convenient in terms of time, money, effort, or comprehension, or because they have achieved standardization by force of long usage or formal agreement.

Following Burt (1) into another analogy, the selection of the factors of descriptive factor analysis may be conventional in the same sense that the selection of Greenwich as the zero meridian was a matter of convention. In contrast, the location of the northern geographic pole is not a matter of convention. By "northern geographic pole" is meant a point or small area with certain properties relative to the earth's rotation. Once so defined the location of the north pole is a question of fact, not of choice.

Thus, by descriptive factor analysis is herein meant the kind of factor

analysis that aims at converting the correlation matrix to a more usable form. The facts the descriptive factorist wishes to deal with are represented solely by the correlation matrix. His factor matrix, like the draftsman's drawing, adds no factual material but contents itself with an exact restatement of the factual material expressed by the correlation matrix. Hence terms like "right," "wrong," "error," "true," "false," are inapplicable to the factors of descriptive factor analysis.

It may be that the precise characterization of descriptive factor analysis is as the kind of factor analysis in which the meaning of the factor matrix is the same as the meaning of the correlation matrix. Different factor solutions for the same correlation matrix actually mean the same; they differ only in being different modes of expression for the same meaning.

Inferential Factor Analysis

Factor studies are frequently found, however, in which the author seems to regard the factor matrix as giving factual material in addition to, but inferred from, the factual material in the correlation matrix. The factors seem to be regarded as observable variables which were indirectly observed and discovered but not directly observed in the study at hand.

Such factorists give the impression that they consider the choice between the various possible factorial solutions to be a matter of right or wrong, truth or falsehood. Indeed, if such factors have objective existence, independent of the particular study, statements concerning the existence of such factors would be true or false. The discovery of an object or process hardly seems to be a matter of convention or arbitrary choice, something true by definition, but seems, instead, to be a matter in which errors can be made.

It may be possible, then, to characterize adequately inferential factor analysis as that form of factor analysis in which the factor matrix has empirical meaning in addition to that expressed by the correlation matrix. Two different factorial solutions for the same correlation matrix could, then, express different and even contradictory meanings. The meaning of the factor matrix which is in addition to that expressed by the correlation matrix is the reference to the unobserved but hypothesized "factors." Since these factors are not referred to by the correlation matrix, the correlation matrix would seem to be serving as the evidence from which the factors are inferred by some process or other. Hence the term "inferential" has been selected to characterize this form of factor analysis.

A Fictitious Inferential Factor Study

The feasibility of the above ambitious program of inferential factor analysis will be doubted by many. It is my opinion, of course, that a good deal of the theoretical discussion in factor analysis represents just such a rejection of the current rather ambiguous formulations of inferential factor

analysis. It may be that this rejection, though understandable, is too hasty. Consider, for instance, the following study which appears to be inferential yet possible.

The problem is the old one of whether there are any functional differences among the retinal cones as, for instance, are hypothesized in the Young-Helmholtz theory. The experimental procedure follows.

The cornea and lens of one eye of an anesthetized animal are removed and a micro-electrode is inserted into the optic nerve. If the technique is delicate enough this may be done so that the electrode penetrates a very few, or even one, of the fibers running from the retina to the cortex.

The eye is now stimulated with monochromatic light of controlled intensity. By increasing the intensity of the light the threshold of the intensity needed to elicit a discharge from the retinal element or elements that have been tapped can be found. This is done for various wavelengths covering the visible spectrum. Since the intensity of the stimulating light is inversely related to the sensitivity of the elements in question it might be desirable, in order to get a measure of sensitivity, to convert the intensities to per cents of the maximum intensity used and take as the measure of sensitivity the complement of each per cent. Sensitivity curves for each position of the electrode could then be plotted.

The score matrix will be formed as follows. The row headings will represent different positions of the electrode, perhaps all in one eye, but more probably several eyes will be used. The column headings will be the various wavelengths used. The entries in each cell will be the sensitivity rating that the electrode in the given position got at the given wavelength. Correlation would be by rows, that is, each sensitivity curve would be correlated against every other and the complete correlation matrix would be formed. This correlation matrix is the complete set of inter-comparisons of each electrode sensitivity curve with every other.

This question now arises: the study is aimed at settling a matter of fact. Are there different kinds of cones distinct from one another as rods differ from cones? Now how is this question of fact to be settled on the basis of these correlations?

Suppose that Thurstone's argument (6, pp. 319-346) is paraphrased as follows. Since the technique was careful and precise it can be hoped that most of the electrode positions tapped only one or two cones. Suppose also that the cones actually do fall into a relatively few and distinct classes, that is, "red" cones, "blue" cones, and so forth. If so, the electrode positions will be "simple" in the sense that most of them tap only one or two kinds of cones while actually three or four or five kinds exist. In such a case the correlation matrix will have a certain pattern, the one called simple structure.

If the above conditions hold for a given study then the correlations will have simple structure. Suppose next that in the study at hand simple structure

does appear. Is it not a plausible inference, then, that this simple structure was due to some such limited number of kinds of cones? If no such structure appears then, of course, the correlation matrix will be regarded as offering no evidence for any such system of kinds of cones. But if the simple structure does appear, the factorist will regard its existence as offering evidence with some degree of plausibility for the existence of a system of cones.

If, then, the structure is of such a character as to render the inference about kinds of cones plausible, the factorist would proceed to rotate the reference axes so as to best represent each of the kinds. The "factors" are thereby transformed from arbitrary reference axes to idealized individuals representative of each of the classes. Using these factors the scores of each wavelength on these factors can be computed, thus leading to the hypothetical sensitivity curves for each factor.

So, from such data, the factorist might attempt to infer the facts that kinds of cones exist, the number of such kinds of cones present, their names ("red," "blue," etc.), and their sensitivity curves.

Granit's Actual Study

This fictitious factor problem is not an invention of my own but is a reinterpretation of a series of experiments by Ragnar Granit (**2**, **3**, **4**). Granit used the micro-electrode, as has already been described, inserting it into the optic nerve of various animals such as frogs, snakes, cats, and guinea pigs. While he did no correlating or factoring, he did determine sensitivity curves for each position of the electrode.

I am not competent to pass on the details of the operative technique, but for the purposes here I would regard criticisms of the operative technique as irrelevant. I would like to take his results at face value simply for the purpose of examining the logic of his conclusions.

Over and over again Granit found sensitivity curves from different positions of the electrode that were practically identical. In fact he found five or six such sets of practically identical sensitivity curves. All sets of curves were narrow, indicating a sensitivity to a restricted range of the spectrum. One set had its peak in the red, one in the yellow, two sets had peaks in the green and two had peaks in the blue. Since over and over again he found practically the same "red" curve occurring, he made the obvious assumption that a specific kind of receptor must have existed to account for the sameness of these curves. In like fashion he assumed a different receptor for each of the different classes of curves. Thus he offers a six-receptor theory of color vision.

It may be obvious now wherein this procedure is essentially the same as with the fictitious factor study. Granit did not actually compute correlations, but in classifying the various sensitivity curves as being practically identical or different he was correlating "by eye" as it were. In factorial

parlance he was noticing that his variables fell into sharply defined clusters when he noted that his curves fell into classes such that, within each class, the curves were virtually identical but that the curves of one class were quite different from those of the other classes.

He then supposed that each cluster was the result of a specific kind of receptor. Since a clustered structure is the special case of simple structure that arises when each variable depends on only one factor, this last assumption seems to be the equivalent of recognizing simple structure and rotating to it.

Finally, having identified the cones, he offers composite sensitivity curves for each type of cone. These curves are arrived at by averaging the appropriate values for the actual curves to get the corresponding value for the composite. Factorially this is closely equivalent to finding the scores of each wavelength on the factors. This is chiefly a matter of getting a sum of the weighted scores of the wavelength for each electrode position. In this case, since the electrode positions were probably very closely clustered with very high and practically equal loadings on the factor, the weights would be virtually equal. Hence the simple average that Granit took would be virtually identical.

The points of factorial interest in Granit's work are not yet exhausted. For one thing, it is interesting to note the analogy with Thomson's sampling theory. The factors that Granit discovers are not single cones, it should be remembered. They are classes of cones and a great many receptors in each retina can be supposed to be in each of these classes. And, of course, Thomson thought that Spearman's g might really be a pool of more fundamental entities. Perhaps implicit in Thomson's thinking was the feeling that a group of things was not a fundamental enough entity. I see, however, no objection to considering a class of objects as a factorial entity and, as a matter of fact, I would suppose that most factors are classes in some sense.

Of more interest, however, are the implications of Granit's study for the higher-order factors. If inferential factor analysis does endeavor to discover and identify entities of some kind, it ought to be general enough to discover entities of as many kinds as possible.

Suppose that in the fictitious factor study it had been found that the clusters themselves were correlated with each other. In such a case some factorists would arbitrarily select an uncorrelated set of factors to represent the clusters. Others, however, would fit the factors to clusters as accurately as possible and would then admit that, as a matter of fact, the factors were correlated and indicative of some higher-order relation among the different types of cones.

What appears in Granit's study? Very frequently a position of the electrode would result in a sensitivity curve much broader than the sensitivity curves of the red, blue, green, and yellow cones. Granit calls this the "dominator" and calls the narrower curves the "modulators." Due to the

resemblance of this curve to the "luminosity curve of the light-adapted human eye" (**2**, p. 12), Granit assumes that this is a brightness receptor. It is not clear, however, what Granit believes this dominator to be. At one point he refers to it as the "dominant receptor" (**2**, p. 13) as if it were another kind of receptor. On the other hand he also thinks of it as a set of modulators:

> The experiments with the cone-eye of the snake suggested that the dominator itself is composed of modulators joined together in such a fashion—either photochemically or by connexions in the retinal synapses—as to operate as a *functional unit* (**2**, p. 13).

Factorially Granit seems to be considering different hypotheses. Let us consider first the notion that the dominator is another kind of cone. In the factor study, then, certain electrode positions would have been present which tapped these receptors. The resulting sensitivity curve would then have overlapped the modulator curves with the result that it would have correlated to some extent with all the curves. This, however, will not produce correlations from one modulator to another modulator. The modulators can remain as an orthogonal set of clusters. The dominator variables will also cluster thickly since their curves are practically identical with one another. This, then, is the description of a structure containing a first-order general factor.

Suppose, however, that the dominator is a grouping of the modulators and does not represent another kind of cone. The sensitivity curve of every electrode position will be contaminated by the fact that it does not represent the modulator sensitivity alone but also, to some extent, the joint action of the rest of the modulators in the dominator of which it is a member. Hence, although the clusters will still appear as before, they will be correlated to some extent since every curve will tend to correlate with every other. Factorially, the dominator is then a second-order factor and no dominator cluster will appear in the first order.

Here Granit is entertaining differing factual hypotheses that would have different factorial results. A factor study could, apparently, lead to a decision between the hypotheses if the factorist let the structure determine the rotations and did not rotate arbitrarily. It might be said that inferential factor analysis intends to settle matters of fact. If so, the procedure cannot entail making before-the-fact assumptions and remain inferential. The existence of second-order factors is not a matter of a priori assumption but something for the empirical evidence to decide. The oft-read remark that "the battery was rotated to simple structure but kept orthogonal due to the greater convenience of an orthogonal frame" seems to represent a confusion of purpose between descriptive and inferential factor analysis.

Meaningfulness of Inferential Factor Analysis

The primary question for inferential factor analysis is, of course, the matter of whether it is even possible. Toward the solution of this problem

we now have the example of Granit's study which seems to have the essential attributes of an inferential study. From correlation-like evidence Granit has made inferences concerning physiological entities which were not directly observed. To me his inference of various kinds of cones from the basis of his evidence seems quite plausible. This, of course, is no argument, but it should be remembered that the original distinction between what we now call rods and cones was probably a similar inference, merely plausible, but one which has proved very fruitful.

It seems clear, furthermore, that Granit's study was not a descriptive factor study. If it were, the factors (that is, the "cones") should be merely ways of expressing the relationship between independent and dependent variables. Granit, however, certainly feels that his cones have empirical meaning in addition to that contained in his observations. He speculates, for instance, as to whether the cones are differentiated by possessing some of the known chemical relatives of visual purple. Such speculations are actually tentative predictions concerning the outcome of further experiments quite different from the original. Granit may be mistaken in thinking that he is justified in such speculations, but if not, then this property of suggesting further and different experiments would indicate that his factors possess meaning additional to that of his evidence.

Still, the commonplace plausibility of Granit's study is not sufficient to dispel the long-standing criticism of inferential factor analysis. I would suggest that this can be done only in the following fashion.

In spite of the great amount of mathematics that surrounds factor analysis, no one has expressed mathematically, or in any other precise fashion, the definition of a factor and its relation to the observational material. In the absence of any such clear expression of what factors are, little of a definitive nature can be said concerning what actually goes on in a factor analysis. This matter of the meaningfulness of inferential factor analysis ought to clear itself up as soon as anyone formulates a precise and adequate language in which the term "factor" is defined and its relation to the rest of the procedures of factor analysis given in detail. This of course is what is known as logical analysis.

How Is the Inference in Factor Analysis to Be Accomplished?

Contingent upon his success in the first problem the inferentialist will then face the problem of offering a theory of inference by which he plans to make his actual inferences. It is one thing to talk about inferring factors from correlations. It is quite another thing to show how it can be done. Since the answer to this problem is not at all obvious, it makes one wonder if the solution is not, in principle, impossible, thus harking back to the first charge of metaphysics.

In Granit's case it was shown, I believe, that certain aspects of the

structure were taken as evidence of some sort of an underlying entity. In fact, Granit was recognizing a special case of what Thurstone would call simple structure, and it is also apparent that Thurstone regards the existence of simple structure as evidence for the operation of underlying unities.

This method, however, has by no means gained universal assent. For one thing the principle is not quite clear. This can be seen quickly enough if one examines the factor plots of various studies. The authors may all talk the same about simple structure, but it will soon be apparent that their understandings must be rather various.

In addition the principle has a rather *ad hoc* quality. It was developed especially for factor analysis. At the same time, however, if one conceives of it as a method of making inferences, it becomes puzzling as to why it is not rationalized on a broader basis. In philosophy, logic, statistics, and mathematics a great deal has been said about how to draw one kind of an inference or another, and it is surprising that inferential factor analysis stands by itself, without reference to any theory of inference. For myself, I share Burt's belief that the inference is meant to be inductive in character, but not his belief that it is impossible.

REFERENCES

1. Burt, C. The factors of the mind. New York: Macmillan Co., 1941.
2. Granit, R. A physiological theory of color perception. *Nature*, 1943, **151**, 11-14.
3. Granit, R. Isolation of the mammalian colour receptors with micro-electrodes. *Nature*, 1945, **155**, 711-713.
4. Granit, R. Sensory mechanisms of the retina. London: Oxford Univ. Press, 1947.
5. Kelley, T. L. Comment on Wilson and Worcester's "Note on factor analysis." *Psychometrika*, 1940, **5**, 117-120.
6. Thurstone, L. L. Multiple-factor analysis. Univ. Chicago Press, 1947.

PSYCHOMETRIKA—VOL. 19, NO. 3
SEPTEMBER, 1954

A SIMPLE PROOF OF A PERSONNEL CLASSIFICATION THEOREM

HUBERT E. BROGDEN

PERSONNEL RESEARCH BRANCH, AGO, DEPARTMENT OF THE ARMY*

A simple algebraic proof of a theorem defining the optimal solution to the personnel classification problem is given. If a set of constants, one for each job, are known, the theorem indicates that each individual should be classified by adding the constants to the estimates of the individual's productivity in the several jobs and selecting the job for which the resulting sum is highest.

In this note, a proof will be advanced for a theorem basic to the solution of the classification problem. This theorem, in somewhat different forms, has appeared before in the work of Brogden (1), Lord (3), Votaw (5), Rao (4), and Dwyer (2). Proofs having varying degrees of rigor have been presented in the publications by these authors. The content of this note offers a reformulation and a much simpler and more easily understood proof.

Definition of Terms

N = the number of individuals.

n = the number of jobs.

C_{ij} = the estimate of individual i's performance in job j. [Elsewhere (1), this author has argued that the C_{ij}'s should be least squares estimates of criteria expressed in terms of a common metric that reflects the value of the individual to the hiring organization. The nature of the C_{ij}'s is, however, irrelevant to the proof to be developed in the present paper.]

Q_j = the quotas; the number of individuals required for each job. The sum of the quotas equals N; that is, all individuals are to be classified.

X = an n by N allocation matrix with elements x_{ij} taking on values of 0 or 1. A numerical value of 1 for a particular x_{ij} has the effect of including the corresponding C_{ij} element in the allocation sum (to be defined), while entries of 0 exclude the corresponding C_{ij} element from the allocation sum. For each i there will be a single x_{ij} having a value of 1 since each individual is placed in a single job. The conditions of the problem demand that the number of entries of 1 for each j corresponds to the quota for that job. Thus, $\sum_i x_{ij} = Q_j$. Use of

*The opinions expressed are those of the author and are not to be construed as reflecting official Department of the Army policy.

X is solely a matter of convenience in symbolizing the selection of elements for the allocation sums. It would have been possible, though awkward, to define verbally the selection of C_{ij} elements for allocation sums.

$\sum_{i,j} C_{ij}\, x_{ij}$ = the allocation sum. From the definition of the allocation matrix, it is evident that summing the cross-products of C_{ij} and x_{ij} is equivalent to summing, across all individuals, the estimated productivities for the job to which each is assigned by a given allocation matrix.

K_j = a set of constants, one for each job. The K_j's are assumed to have numerical values such that if individuals are allocated to the job in which their value of $(C_{ij} + K_j)$ is highest, the number allocated to each job will correspond to the number specified by the quota for that job.

X' = a particular allocation matrix with elements x'_{ij} that allocates individuals to jobs according to the highest value of $(C_{ij} + K_j)$. Stated otherwise, X' has elements which take on the value of one for each i when the value of $(C_{ij} + K_j)$ is optimum for the given i. Since X' is a particular X, it conforms to other limitations placed on X. In particular, $\sum_i x'_{ij} = Q_j$.

The Problem

The classification problem may be stated as follows: given estimates of productivity (C_{ij}'s) for each of N individuals in n jobs and given the number of individuals to be placed in each job (Q_j's), what allocation of men to jobs will yield a maximum allocation sum? We might also ask: how should X be defined so that $\sum_{i,j} C_{ij}\, x_{ij}$ will be at a maximum?

The Proof

We wish to prove the following theorem: if appropriate values of a set of constants, $K_1 , \cdots , K_n$, exist and the quantities $(C_{ij} + K_j)$ are available for each man in each job, an allocation of men to jobs by selecting for each man that job for which the quantity $(C_{ij} + K_j)$ is highest will yield an optimal solution; that is, this procedure will result in an allocation sum that is optimal for the given quotas. Stated otherwise: the allocation sum associated with X' is equal to or greater than the allocation sum associated with any other allocation matrix; that is, $\sum_{i,j} C_{ij}\, x'_{ij}$ is equal to or greater than $\sum_{i,j} C_{ij}\, x_{ij}$.

Since X' allocates each man to the job in which the value of $(C_{ij} + K_j)$ is highest we know that

$$\sum_{i,j} (C_{ij} + K_j)x'_{ij} \geq \sum_{i,j} (C_{ij} + K_{ij})x_{ij} , \tag{1}$$

or

$$\sum_{i,j} C_{ij}x'_{ij} + \sum_{i,j} K_j x'_{ij} \geq \sum_{i,j} C_{ij}x_{ij} + \sum_{i,j} K_j x_{ij} . \tag{2}$$

Now

$$\sum_{i,j} K_j x'_{ij} = \sum_j (K_j \sum_i x'_{ij}), \tag{3}$$

and similarly

$$\sum_{i,j} K_j x_{ij} = \sum_j (K_j \sum_i x_{ij}). \tag{4}$$

Since X' and X have, by definition, Q_j entries of 1 and $n - Q_j$ entries of 0, we have, for any given j

$$\sum_i x'_{ij} = \sum_i x_{ij} = Q_j . \tag{5}$$

Hence,

$$\sum_j (K_j \sum_i x'_{ij}) = \sum_j (K_j \sum_i x_{ij}) = \sum_j K_j Q_j . \tag{6}$$

Substituting in equation (2) we have

$$\sum_{i,j} C_{ij} x'_{ij} + \sum_j K_j Q_j \geq \sum_{i,j} C_{ij} x_{ij} + \sum_j K_j Q_j , \tag{7}$$

or

$$\sum_{i,j} C_{ij} x'_{ij} \geq \sum_{i,j} C_{ij} x_{ij} , \tag{8}$$

which is what we set out to prove.

If a reject group is involved, the solution we have considered will apply if a dummy job is set up and the estimates of productivity for this dummy job are set at zero for all N cases. Use of such a job implies, as is obviously the case, that individuals who are rejected (placed in the dummy job) contribute zero to over-all productivity.

Discussion

In the foregoing proof, the existence of an appropriate set of K_j's was assumed. Some further comment on this assumption is appropriate here. First, it should be noted that, in use, the validity of this assumption is checked empirically. If K_j's are found such that classification by the highest value of $(C_{ij} + K_j)$ gives the proper number of cases in each job then the K_j's are shown to exist for that example. Thus, invalidity of this assumption cannot lead to an incorrect conclusion. In the writer's experience with empirical applications, the appropriate K_j's have always been found to exist.

A number of methods for determining the K_j's have been proposed (**1**, **2**). Generally speaking, these methods are iterative and involve raising the values of K_j for jobs in which shortages exist until by successive tries a fit to quotas is obtained. It has been argued that proof of the existence of the K_j's is necessary in showing the utility of such iterative procedures. The

following comments bear upon this problem, although a formal proof will not be attempted. If the range of C_{ij}'s for all jobs is finite and the K_j's for all but one job are fixed and finite, variation of K_j for the remaining job can obviously result in variation of the number placed in this given job from zero to N. With each increase in K_j , the number placed in the given job will either increase or remain stationary. In theory, if the estimates of productivity are perfectly continuous, if the multi-dimensional frequency solid is not discontinuous, and if the number of cases is infinite, infinitesimal increments in K_j will produce infinitesimal increments in the proportion of cases assigned to the given job. Hence, K_j for the given job exists no matter what finite values of K_j are chosen for each of the remaining jobs. The same reasoning applies to the K_j's for all jobs. Hence, we can generalize our statement to say that all K_j's exist.

The statements of the preceding paragraphs are subject to some qualifications if the K_j's and C_{ij}'s are discontinuous variables. In that event, tied scores will be found and an exact fit to quotas may not be obtained merely by variation of the K_j's. With tied scores there are alternate optimal solutions and an exact fit to quotas is achieved by arbitrary placement of the individuals involved.

REFERENCES

1. Brogden, H. E. An approach to the problem of differential prediction. *Psychometrika*, 1946, **11**, 139–154.
2. Dwyer, P. S. Solution of the personnel classification problem with the method of optimal regions. *Psychometrika*, 1954, **19**, 11-26.
3. Lord, F. M. Notes on a problem of multiple classification. *Psychometrika*, 1952, **17**, 297-304.
4. Rao, C. R. Advanced statistical methods in biometric research. New York: Wiley, 1952.
5. Votaw, D. F. Methods of solving some personnel classification problems. *Psychometrika*, 1952, **17**, 255-266.

Manuscript received 9/18/53
Revised manuscript received 11/3/53

PSYCHOMETRIKA—VOL. 19, NO. 3
SEPTEMBER, 1954

VERIFICATION OF HYPOTHESIZED FACTORS IN ONE HUNDRED AND FIFTEEN OBJECTIVE PERSONALITY TEST DESIGNS

R. B. CATTELL, S. S. DUBIN, AND D. R. SAUNDERS

UNIVERSITY OF ILLINOIS

This paper reports on the third of a series of four experiments using similar concepts and methods for objective personality measurement, and overlapping test batteries. One hundred students were measured with 115 tests. The scores were correlated and factored by a re-iterated multiple group centroid method. The 17 factors thus obtained were rotated toward a clear simple structure. The relation of the rotated factors to earlier ones is indicated, but no extensive interpretation is attempted.

1. *Principles Underlying the Present Research*

This study is the third in a series of four related factorial studies of objective personality tests carried out between 1944 and 1951 (**3, 8, 11**). Since the fourth study employs psychotic subjects (**11**), this report concludes the "normal" series. The following methods and objectives apply to all of these studies:

a. The tests are based upon hypotheses derived from earlier factorial studies of behavior ratings (**2, 5, 6**), or questionnaire responses (**7**). It is assumed that the factors which have been found in these realms can also be observed as factors in the realm of objective tests. It is further assumed, in view of the previous use of the personality sphere concept (**2**) for the selection of rating variables, that an adequate sampling of objective tests will be obtained by using several intended to measure each of the rating and questionnaire factors.

b. The tests are objective in two senses. They measure behavioral responses without accepting verbal symbols at face value. They are independent of the judgment of the individual examiner in scoring. Thus, the new tests differ both from "personal inventories" and from such tests as the Rorschach, Szondi, TAT, etc.

c. The four successive factorial studies carry forward into each experiment a sufficient number of "marker variables" to define the factors found in earlier studies. The first study (**3**) used a battery containing markers for the sixteen or so factors that had appeared in the literature up to 1944 (**2**); its aim was to confirm and consolidate the previous work. The second (**8**) and third studies have brought entirely new test designs into the factorial reference frame, in order to expand our knowledge of existing factors and to

generate new common factor space important to our understanding of personality. The third and fourth (**11**) studies introduce a new direction of expansion into the program, for they aim to compare the factor structure of normal and abnormal populations on substantially the same battery—and thus the fourth study introduced only ten new tests.

d. Emphasis has been given to psychological ingenuity and the invention of new objective personality tests rather than to precise determinations of factor loadings. It is felt that a fruitful exploration of "personality space" will depend more upon recognition of factor *patterns* than upon particular tests with high factor saturation. It will require a particularly varied set of tests to define all the known or suspected personality factors with even sufficient loading for recognition, and to set them apart from ability factors. Therefore, it must be stressed that our tests are deliberately brief, and that our reliabilities and loadings are not expected to be so high as for longer tests.

Of course, knowing the reliability, it is possible to estimate the chances of obtaining useful factor validity by lengthening a given test. It is intended that the work reported here will be followed by a new phase of confirmation and consolidation of factor information, preparatory to selection, lengthening and refinement of tests holding promise for further research or application.

2. *Description of Tests and Their Administration*

The 115 tests listed in Table 1 include 80 group tests and 35 individual ones (indicated by * in Table 1). The group tests were given in a constant order during several sessions to over 100 male undergraduate students, while the individual tests were given in an uncontrolled order during the same two-month period. Thus, as is usual in *R*-technique studies, function fluctuation and test-sequential effects will contribute to the error variance of measurement.

In addition to the usual serial numbers of the variables, which refer to the accompanying matrices, the Master Index Numbers are also given, to facilitate reference to the same variables throughout the related studies (**3, 8, 11**) and elsewhere. (The writers are indebted to Dr. Walter Gruen for his care in the upkeep of the "Encyclopedia of Personality Tests," containing the tests used during this decade of research in the Laboratory of Personality Assessment. All our published researches report, in addition to the matrix numbers of the particular sample of tests, the "Master Index Numbers" of this Encyclopedia.) When improvements could be made in administration of marker tests without changing their essential process, this was done, and the improved test was given a new Master Index Number along with a cross-reference under its old number. Thus, the following information about each variable is given in the Master Index column of Table 1: first, reference to (**3**) and/or (**8**) (as F_1 and F_2) if the variable is a

marker from either or both of these studies; second, its Master Index Number for the form used in (3) and/or (8); third, its Master Index Number for the form used here. Both the first and/or second parts will be omitted if they do not apply.

Adjectives such as *fast, few,* etc. are included in the test title whenever there might otherwise be doubt as to the positive direction of scoring. Whenever any of the tests would normally have been given without time-limits, the number of "right" responses was divided by the number of items completed.

Only the utmost brevity of description for the 115 variables is possible in Table 1*. The rationale of a given test has been limited to a statement indicating the behavior rating factor it was expected to measure, since the hypothetical nature and empirical associations of these factors are discussed elsewhere (2, 5, 6). In the case of marker variables—the first 47 variables listed—reference is made to their previous description (3, 8, 10). Corrected split-half reliabilities are shown in Table 2 for all variables on which they can be computed.

It will be noticed that any newly-tried score, e.g., effect of approval or disapproval on performance, tendency to agree, or discrepancy of aspiration, is found for at least two tests employing different material. This is to permit an examination of whether the supposed formal process exists as such. These duplications, as well as the practice of including more than one score dependent upon the same tests, so as to run the risk of correlated errors, can tend to produce narrow common factors peculiar to the tests involved. It is gratifying that no such artifacts were found in this study, perhaps because of the large magnitude of the variance required for recognition of a factor in this research.

3. *The Factor Analysis*

Exactly 100 subjects were chosen whose test scores were most complete and free from extraneous effects such as illness, absence, or misunderstood instructions. For each test the raw scores were ranked; tied scores were assigned the ranks tied for at random. The ranks were then converted into a "normal" distribution with mean of 10 and range from 1 to 19. The normalized scores were correlated by the product-moment formula. [These operations were performed in IBM cards by (1) punching a card for each subject, (2) shuffling the deck thoroughly, (3) sorting on raw score, and (4) reproducing an ordered set of normalized scores into the deck. The 602-A was used to find $\sum xy$ for five pairs of variables at a time, and these were converted to r's by a single multiplication and subtraction.]

*The Institute for Personality and Ability Testing, 1608 Coronado Drive, Champaign, Illinois, is planning to publish these tests and their instruction booklets, during 1954, beginning with the tests of highest factor loading.

TABLE 1

List of Objective Personality Test Variables

Serial No.	Master Index No.	Description
1	2b(170)	Lack of Rigidity (Hidden pictures: a priori key). In (10). Factor C (6).
2	2c(171)	Lack of Rigidity (Hidden pictures: flexible key). Any alleged hidden picture accepted. In (10). Factor C (6).
3	2	Lack of Rigidity (Hidden words). Letters of words given in random order. 40 words in four minutes. In (10). Factor C (6).
4	2(IV)(i)(172) F_1F_2	Motor Rigidity. Letters written with backward movement. Perseveration Test 'p' in (4, p. 234). Factor G- (6).
5	2(IV)(ii)(173) F_1F_2	Motor Rigidity. Personal name written backward. Perseveration Test 'p' in (4, p. 234). Factor G- (6).
6	2(IV)(iii)(174) F_1F_2	Motor Rigidity. Numbers written backwards. Perseveration Test 'p' in (4, p. 234). Factor G- (6).
7	3b(127) F_1F_2	Speed of Judgment. "Working quickly" on a variety of opinion statements (not intelligence items). Factor F.
8	*4(I)(175) F_1F_2	Slow Reaction Time (Regular signal). In (3). Factor A- and Thurstone's perceptual factor (13).
9	*4(II)(176) F_1F_2	Slow Reaction Time (Irregular signal). In (3). Factor A- and Thurstone's perceptual factor (13).
10	*5 F_1F_2	Ratio Regular to Irregular Reaction Time. Variable 8 divided by variable 9. In (3). Factor A-.
11	*6c(177) F_1	Motor Speed. Sum of speed scores on variable 39, corrected for coordination, and 5 writing speed scores on variables 4, 5, and 6.
12	7 F_1	Speed of Perceptual Closure. 12 incomplete (Gestalt) pictures and 12 incomplete words. Score equals number completed in 70 seconds. In (3). Represents Thurstone's perceptual factor.
13	*8 (F_1F_2)	Rate of Alternating Cube Perspective. In (8).
14	*9c(178) F_1F_2	Slow Tempo. Time to make drawing, answer questions, measure lines, draw squares, and produce ink-blot associations. S not instructed to work rapidly, but to go at normal rate. To mark factor in (3).
15	11b(130) F_1F_2	Intelligence (Verbal-Numerical Ability). Primary Mental Abilities sub-tests. Factor P.
16	13 F_1F_2	Oscillation. Variation of successive 7-second performances on cancellation, checking, etc. In (3).
17	15 F_1	Cursive Miniature Situation (Excess Circles). Defined in (3). Group form.
18	21 F_1F_2	Low Questionable Reading Preferences. Defined in (3), and slightly changed by item analysis. Factor G or M.
19	22b(131) F_1F_2	Ratio Chance to Purposeful Observation and Memory. Defined in (3). Slightly modified, and scored here in opposite sense. Factor H.
20	24a(i)(179) F_1	Ratio Initial to Later Performance. Ratio of first to third minute in observing, counting, and recording. Factor G.
21	24a(ii)(180)	Ratio Initial to Later Performance (Backwards Reading). Same as variable 20 on task of reading sentences printed backwards.
22	25(F_1)	Ratio Emotional to Non-Emotional Recall. In (3). Factor C.
23	25c(132)	Ratio Emotionally-Interesting to Emotionally-dry Recall. Uses stories rather than a word list. Ideas recalled in strongly emotional stories divided by words recalled following lecture-type material. Factor G-.
24	26 F_1	Ability to Handle Surprises (Riddles). In (3) and (7). Factors B and negative of "ideational rigidity" in (7).
25	28 F_1	Dynamic Momentum. Time spent on individual anagrams and problems, instead of switching to new problem stems. Factor G. In (3).
26	29c(133)	Criticalness of Judgment. Quality of judgment of human performances which could be graded by S on a "hostility-geniality" dimension. In (3). Factor A- or E.
27	30b(134) F_1	Ratio Self-criticism to Criticism of Others. As in (3), except that "approvals" not included in ratio.

TABLE 1 (continued)

List of Objective Personality Test Variables

Serial No.	Master Index No.	Description
28	31(i)(181) F_1	Fluctuation of Attitudes 1. (Variable 34 material). Sum of absolute values of shifts on re-presentation of attitude statements. In (3). Factors A and C-.
29	31(ii)(182) F_1	Fluctuation of Attitudes 2. (Variable 33 material). Scored like variable 28.
30	31(ii)(183) F_1	Fluctuation of Attitudes 3. (Variable 79 material). Scored like variable 28.
31	31(iv)(184) F_1	Fluctuation of Attitudes 4. (Variable 78 material). Scored like variable 28.
32	33 F_1	Myokinesis (Mean Length). Mean size of twenty attempts to reproduce a 1-inch line while blindfolded. In (3). Factor F.
33	34 F_1F_2	Immaturity of Opinion. Change in opinion upon re-presentation of attitude statements in direction indicated by newly presented facts, which should have been thought of earlier. In (3) and (8).
34	35 F_1F_2	Suggestibility to Authority. Change in opinion upon re-presentation of attitude statements as quotations by authorities. In (3) and (8), but slightly modified. Factor E-.
35	36 F_1F_2	Ability to State Assumptions. Based on Watson-Glaser test. In (3) and (8). Factor I.
36	38b(135) F_1F_2	Ratio Consonant to Dissonant Opinions Recalled. In (3) and (8). Factor H-.
37	38c(185)	Ratio Consonant to Dissonant Recall (Corrected). Similar to variable 36, but each number recalled is first divided by the number of like-valued statements checked.
38	39b(136) F_1F_2	Ratio Color to Form in Sorting. A sorting task in which neither color nor form is an entirely adequate guide. In (3) and (8), but modified for group presentation using slides.
39	*41 F_1F_2	Two-Hand Coordination. In (3) and (8). Marks Thurstone's perceptual factor (13).
40	*42 F_1F_2	Body Sway Suggestibility. As in (3), (8) and (11). Factor C- (12).
41	*43 F_1F_2	Magnitude of Mean PGR Deflection. Mean percentage conductance. In (3) and (8). Factor H.
42	*43(i)	Ratio PGR Response Mental to Physical Stimuli. Mean conductance change for 7 mental stimuli (e.g., shock). These means were added to obtain variable 41.
43	*46c(137) F_1	Impairment of Maze Performance by Shock. S made 4 runs of a blind pencil maze, with instructions to get as far as possible. During runs 2 and 3 he was given a shock, but was not told that this would be punishment for poor performance. Score is ratio of total distance in runs 1 and 4 to total of runs 2 and 3. Factor H.
44	*62b(138) F_2	Ratio Emotionally-Interesting to Emotionally-Dry Reading Rate. (cf. variable 23). Ratio reading speed on four passages--2 gripping and 2 dull, when instructed to read quickly. Factor G-.
45	65b(139) F_2	Logical Consistency of Attitudes. Improved from (8), but still not incontestably logical. Factor C.
46	67b(141) F_2	Extremity of Viewpoint, 1. (cf. variables 85 and 96). Ratio of "absolutely true" and "absurd" checks to more moderate alternatives on 100 attitude statements. In (8).
47	*83 F_2	Fidgetometer Frequency. Number per minute during middle part of individual test session. Best with record of body as well as foot movements. Factor 6 in (8).
48	*91	Autistic Suspicion. Frequency with which S states people are pointing at him, while under sheet in roomful of spectators. Factor L.
49	*92	Low Readiness to Imitate Animal Sounds. S instructed to "make a noise like a cat," etc. Score is mean time to respond to instruction. Factor F-.
50	*93	Low Effect of Restraint on Performance. S counted number of shapes within complex figures (a) normally and (b) when clamped within physical restraints. Factor C.
51	*93(186)	High Effect of Restraint on Memory. S recalled 8 names of well-known people under each condition of variable 50. Ratio taken in the opposite direction, however.

TABLE 1 (continued)

List of Objective Personality Test Variables

Serial No.	Master Index No.	Description
52	*94	Mean Flicker Fusion Frequency. Mean of 10 up and down runs using Strobotac, under standard distance and illumination. Factor of "rigidity" in (9).
53	*95	Decrease Myokinetic Movement under Threat. Repetition of variable 32, with shock "punishment" for 2 predetermined poor lines out of 10.
54	*96a(i)(187)	Low Reduction of Reading Speed by Frustrated Preference. S starts to read 2 stories and indicates a preference to continue with one. He is then given the other. Ratio of new speed to previous speed is scored. Factor C.
55	*96a(ii)(188)	Low Reduction of Computing Speed by Frustrated Preference. Ratio of computing speeds before and after administration of test 54.
56	*97	Increase in Length of Time Interval Perceived during Ergograph Work. Ratio of S's estimate of 2 time intervals while pulling hard to actual time taken to fatigue S on ergograph task.
57	98	Cursive Miniature Situation (Proportion Correct). (cf. variable 17). Scored as right minus wrong decisions per minute.
58	98b(143)	Cursive Miniature Situation (Number Decisions). (cf. variable 17). Scored for absolute number of correct decisions.
59	99	Ratio Verbal to Numerical Ability. (cf. variable 15). Ratio of PMA sub-tests. Factor C-.
60	100	Optimism over doing Good. S rates the feasibility of 30 "reforms" in a test of "judgment of human nature." Factor A.
61	101	Aspiration Level in Coding. Algebraic sum of five differences between first-page performance and second-page aspiration, in terms of numbers of items completed.
62	102	Tendency to Consider Remarks "Obvious." S checks which of 20 items may be considered "obvious, commonplace, and scarcely worth saying."
63	103(i)(191)	Low Amount Considered Possible in Given Time for Others. S estimates how much time required to complete 20 various tasks for others (multiple-choice).
63	(cont.)	Large estimates lead to high scores. Factor F-.
64	103(ii)(192)	Low Amount Considered Possible in Given Time for Self. Like variable 63, except that tasks are to be done by S. Factors E and F.
65	104	High Self-Estimate of Personal Prowess. S checks degree to which he possesses 30 socially esteemed items.
66	105	Tendency to Perceive Threatening Objects. S presented with 16 roughly sketched collections of objects to be identified. Score is proportion of perceived objects which are dangerous (e.g., weapons, storms, pitfalls) rather than benign.
67	106	Unreflective Acceptance of Unqualified Statements. S checks which of 39 statements, that are so ambiguous as to require qualification, he will endorse as "absolutely certain." Factors F, G-, and K-.
68	106b(144)	Agreement with Platitudinous Statements. S indicates his degree of agreement with 30 platitudes. Factor G-.
69	107	Knowledge of Etiquette. 24 multiple-choice items. Factor K.
70	108	Self-confidence on Untried Performances. This is based on a second part of test 93. S indicates how well he would expect to master each of a variety of accomplishments he says he has not tried, given practice and opportunity. Factor E.
71	109	Ratio Pleasant to Unpleasant Associations. S lists in separate 2-minute periods (a) unpleasant past, (b) pleasant past, (c) unpleasant future, and (d) pleasant future associations. Score equals b/a plus d/c. Factor F.
72	110	Ratio Anteversion to Retroversion. Based on same performances as variable 71. Score is c/a plus d/b. Factor F.
73	111	Excess of Adverse over Favorable Self-Reference in Events. 1. S checks on a list of 30 social, political, and economic movements which are likely to affect him (a) favorably, (b) unfavorably, or (c) not at all. Score equals b minus a. Factor L.

TABLE 1 (continued)

List of Objective Personality Test Variables

Serial No.	Master Index No.	Description
74	111(i)	Amount of Self-Reference in Events. Based on same performance as variable 73. Score equals a/c plus b/c. Factors E and L.
75	112	Ratio of Adverse over Favorable Self-Reference in Events, 2. Based on same performance as variable 73. Score equals b/a.
76	113	Little Breakdown of Reality Principle. S checks a wide range of beliefs scaled on a reality versus autism dimension. The latter extreme includes, e.g., pleasure principle, wishful thinking, irrational fears, ego defense distortions, etc. Factor C.
77	114	Underestimation of Performance. Sum of differences between performance on a page of problems and prior estimates by S. Factor E-.
78	115(i)(193)	Ego Weakness (Avoidance of "Unsuccessful" Responses). (cf. variable 31). Similar to tests 33 and 34. Scored for S's tendency to shift away from designated "unsuccessful" responses upon re-presentation of materials. Factors G- and C-.
79	115(ii)(194)	Ego Weakness (Avoidance of "Neurotic" Responses). (cf. variable 30). Like variable 78, except that certain responses were designated "neurotic." Factors C- and G-.
80	116	Cynicism, Hard-Headed Judgment. S indicates agreement with attitudes toward social issues which are presumed to be hard-headed and cynical as opposed to sentimental and wishful in thinking. Factors A- and L.
81	117	Social Good Taste. A test of information about upper-class tastes in furnishing, eating, reading, etc. Factor K.
82	118	Sense of Poetic Aptness. A multiple-choice test of ability to choose good last lines for poems, keyed using the original rare sources. Factors I and possibly K.
83	119	Artistic Taste in Color-Blending. Number right among 30 paired comparisons of color combinations, as keyed by artists, divided by number of choices made. Factors I and possibly K.
84	*120	Ratio Inaccuracy to Speed. Sum of errors over number-dons fractions for variables 12, 16, 20, 39 and 43.
85	121	Indecisiveness of Viewpoint. (cf. variable 46). Number of middle judgments made on 5-point scale for 100 attitude statements. (Variables 46 and 85 are based on non-overlapping groups of items, based on an odd-even split, to eliminate an otherwise high spurious correlation between them.)
86	*122	Static Ataxia. Measured by body sway, as in variable 40, but without suggestion. Factor C-.
87	*123	Ratio Working to Relaxed Time-Estimates. S estimates length of 2 time intervals while lifting a heavy weight to the point of fatigue, and 2 25-second intervals while doing nothing. Ratio of former to latter.
88	124b(195)	Tendency to Agree with Optimistic Statements. Based on the same performance as variable 60. Score equals ratio agrees to disagrees considering only statements answered optimistically by S.
89	125	Ratio Personal to Institutional Values. S indicates degree of agreement with items presumed to distinguish adherence to internal, conscience values rather than social conventions and habits. Factors I and K.
90	126	Ratio of Male to Female Person Choices. (a) "Artistic preferences of S among male and female sculptures. (b) "Discussion" preferences among male and female companions. Score equals total male choices over total female choices. Factor of masculinity in (8).
91	145	Adaptability of Aspiration Level in Coding. Based on the same performances as variable 61. If a is an initial estimate of performance, b the immediately following actual performance, and c the next estimate of future performance, the score equals the sum of (c minus a) over the sum of (b minus a), for eight repetitions. Factor C.
92	146	Accuracy in Perceptual Closure. Based on the same data as variable 12. Score equals number correct divided by number completed. Factor F.
93	147	Estimated Breadth of Experience and Accomplishments. S indicates his degree of familiarity and skill with 30 various experiences, hobbies, etc. Factor K.
94	148	Aspiration for Improvement in Coding. Score is number of tasks completed on a first page minus S's subsequent estimate of his next performance. (cf. variable 77).

TABLE 1 (continued)

List of Objective Personality Test Variables

Serial No.	Master Index No.	Description
95	149	Reduction of Aspiration Level by Failure. Based on the same performances as variables 61 and 91, but not 77. Score is sum of differences between estimates of next performance and immediately preceding performance. This differs from variable 77 in that there is immediate past experience of performance, and from variable 94 in that there has been opportunity for an estimate based on performance to be compared with performance. Factor C-.
96	150	Extremity of Viewpoint. 2. (cf. variable 46). Same except for use of material from test 62. Factor C.
97	150b(196)	Agreement with Obvious Remarks. Based on same material as variable 62. Score is number of items agreed with or partly agreed with divided by number of items disagreed with or partly disagreed with.
98	151	Ratio of "on-Principle" to "in-Particular" Decision Time. Items in which the decision is between two generalizations are compared with items in which decision is between particular, factual possibilities. Factors K and possibly C.
99	152	Tendency to Agree. Based on same performance as variable 60. Score is number of agreements minus number of disagreements.
100	153	Cursive Miniature Situation (Speed). (cf. variables 17, 57, and 58). Scored for total number of decisions, both right and wrong.
101	154	Cursive Miniature Situation (Slanting Lines). This is a count of slanting lines crossed and measures failure to follow instructions. This is known to be higher in delinquents (4). Factors C- and G-.
102	*155	Range of Flicker Fusion. Difference between largest and smallest numbers recorded for flicker fusion frequency.
103	*156	Lag of Flicker Fusion. Difference between mean of 4 upward runs and mean of 4 downward runs to determine flicker fusion frequency. Possible factor H-.
104	157	Number of Jokes Considered Funny. Total points from ratings of S's reactions to jokes told on a phonograph record. 1 point for negative reaction, 2 points for none or neutral response, 3 points for smile, and 4 points for laughter.
105	*158	Increase of Tempo under Threat. Increase of speed due to shock, as measured in connection with tests 32 and 53.
106	159	Scatter in Other-Referent Time Estimates. Based on same performance as variable 63. Score is sum of deviations about mean estimate when multiple-choice answers are numbered in order of duration.
107	160	Scatter in Self-Referent Time Estimates. Based on same performance as variable 64. Computed analogously to variable 106.
108	*161	Ratio of Approval to Disapproval Cancellation Performances. Comparison of speed and accuracy when examiner said "Good, good" with same when he said "You are doing poorly." Factor E-.
109	*162	Ratio of Normal to Shock Cancellation Performance. Comparison of a normal 30-second period with one in which shock is administered regardless of performance. Factor I.
110	*163	Ratio of Approval to Disapproval Word Closure Performance. Like variable 108, but using material of test 12. Factor E-.
111	*164	Ratio of Normal to Shock Word Closure Performance. Like variable 109, but using material of test 12. Factor I.
112	*165	Ratio of Approval to Disapproval Classification Performance. Like variable 108, on task requiring S to select one word from 5 which does not fit the concept established by the other 4. Factor E-.
113	*166	Ratio of Normal to Shock Classification Performance. Like variable 109, but using material like test 112. Factor I.
114	167	Immediate Memory Efficiency. S's total recall in immediate recognition and reproduction of attitude material contained in various tests. Factors B and perhaps H.
115	169	Awareness of Errors Made. S is instructed how to respond with wrong names to colors, or with learned paired word-associates. Score is number of errors corrected by S divided by total made. Factor G.
116		Random Variable. 1. Two random variables were included in the correlation matrix, in order to provide a basis for determining the level of chance loadings.
117		Random Variable. 2.

Eighteen clusters were selected from the correlation matrix and used as the basis of a multiple group centroid factor analysis, according to Thurston's method (**9, 14**). The factorization was repeated three times, starting each iteration with the communalities obtained from the previous one. After four iterations the communalities were apparently stable*. One factor was lost because of linear dependence of the original clusters.

Rotation was begun with the results of the third iteration, but the transformation matrix obtained after several rotations was applied to the results of the fourth iteration and the rotation toward simple structure continued. When the proportion of variables in the ± .10 hyperplanes became stable, the rotation was stopped. The simple structure can undoubtedly be made more clear by small rotations, but there is little likelihood that any important changes in interpretation would then be called for.

Owing to the sizes of the matrices involved, only the final rotated matrix (Table 2) and matrix of correlations between reference vectors (Table 3) are shown here. [The correlation matrix, unrotated factor matrix and transformation matrix may be obtained from the American Documentation Institute, c/o Library of Congress, Washington 25, D.C., by ordering document 3629 and remitting $2.00 for microfilm (images 1 inch high on standard 35 mm film) or $2.00 for photocopies readable without optical aid.] The communalities in the last column of Table 2 are based on the unrotated factor matrix. The factors have been presented in order of decreasing contribution to total variance, since this information is of *some* help in matching factors in other studies. (Ideally, for matching, this should be based only on the common variables.) It will be noted that the mean number of variables within a ± .10 hyperplane is 71, or 61 per cent, which approximates the criterion of past experience that calls for 2/3 of the variables with such data. Factors 3 and 6 are the poorest, but defy improvement of their hyperplanes.

4. *Factor Interpretation*

In each of the Tables 4-20 we have set out:

a. The half dozen or so highest loaded variables shown in Table 2, stopping at a natural cutting point. The verbal titles of Table 1 have been reflected to agree with the sign of the loading.

b. Occasional variables with a lower loading. With respect to such variables, it should be noted that the highest loading and communality for either of two random variables were 0.31 and 0.24, respectively. (For 100

*It will be noted that some communalities are in excess of the corresponding estimates of reliability. There is, of course, a margin of error in both estimates. In addition, communality estimates including the contribution of chance loadings on a number of factors should be expected to exceed reliability estimates based on the assumption of a single common factor.

TABLE 2

Rotated Factor Matrix

Matrix No. of Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	h^2	Reliabilities
1	21	01	10	-47	-09	02	13	00	-10	02	-05	36	16	-25	-18	04	05	57	.68
2	08	-02	00	-40	-26	-09	22	01	-03	-06	-15	57	04	-20	-08	35	02	72	.75
3	-17	-43	10	15	19	23	-01	-14	09	-05	-22	09	05	-10	-02	06	-03	63	.97
4	-14	-04	-23	04	00	02	-65	-06	-08	-00	-18	-06	-02	-17	03	16	09	62	
5	-13	-05	-08	-10	00	-08	-61	-00	-04	08	05	03	12	05	16	-07	-04	63	
6	-15	-05	-17	06	00	-38	-48	06	-06	04	-01	-01	-07	12	-06	-18	06	47	
7	14	08	08	-04	02	32	30	09	10	09	-13	09	-23	-14	20	-13	-10	38	
8	-20	34	19	-13	31	20	-21	08	10	02	08	-09	-06	-14	-61	08	-09	71	
9	-16	47	24	-04	-26	19	-04	21	12	-07	-05	-13	-14	-20	-48	17	-09	69	
10	04	-21	-16	-05	23	-01	-08	-10	-05	08	18	-10	03	06	-10	-13	-03	23	
11	06	-14	-09	-10	15	21	-01	-19	06	05	-03	-13	-08	-03	03	13	02	22	
12	10	13	15	-03	08	04	19	-03	-01	-09	-20	01	20	-07	-11	58	05	53	
13	02	-08	18	17	26	14	-21	00	02	04	-20	32	-02	-02	20	16	31	52	.85
14	-18	-21	07	-08	16	-43	02	31	02	13	13	-03	-03	25	-19	-09	-11	45	
15	03	-09	40	03	-07	31	04	-19	36	01	-05	08	-23	11	-11	-00	-02	52	
16	-02	-30	08	21	12	-17	-03	-09	09	31	-13	10	-31	25	18	-04	27	55	.51
17	08	-22	-09	-04	-04	-20	02	04	-07	07	24	-11	-06	07	26	-10	04	26	
18	-01	12	30	03	-22	-25	01	31	-12	01	23	-02	-26	19	18	18	01	53	
19	-00	08	07	15	07	-03	04	07	03	05	-09	01	00	06	08	15	23	15	
20	09	-14	07	02	-05	04	16	-15	14	-26	16	-27	-01	-05	08	-15	-08	22	
21	07	03	03	17	-03	-15	09	-06	-18	01	02	12	06	35	-04	07	21	35	
22	07	15	03	15	06	30	02	05	13	-33	00	13	02	-28	-02	11	-03	34	
23	00	00	02	09	-00	21	04	06	08	02	-05	15	02	-33	18	27	-26	37	.35
24	12	05	47	-01	01	-07	-04	-05	-00	-10	-14	00	-12	14	-38	-07	15	53	
25	22	12	21	-08	-09	-07	12	08	00	-13	-02	-08	16	07	-05	03	04	23	
26	-12	04	23	-01	12	-06	-02	02	21	03	22	03	15	57	01	-02	-02	58	.59
27	-07	04	12	01	-38	01	-15	11	06	15	11	20	09	-06	-08	21	17	37	.85
28	-67	07	-16	-05	07	03	-05	-06	13	06	-06	03	12	06	-03	-03	-03	69	.69
29	-53	27	-03	01	05	-13	-09	-13	-01	-06	-18	-03	-03	13	-12	12	05	54	.45
30	-58	20	-13	10	06	12	-15	-08	03	01	02	08	02	-01	-08	07	05	59	.25
31	-59	00	-18	-16	-02	-08	-08	05	-03	07	06	00	20	01	-11	03	-09	58	.68
32	02	07	-08	10	06	06	-03	12	-26	23	26	07	25	-08	-20	-02	08	34	
33	-07	12	-05	-56	-25	07	-16	06	24	-07	-02	-10	16	-16	03	16	00	67	.37
34	-22	29	-10	07	07	08	04	-26	06	-05	01	-27	08	04	07	10	-03	39	.55
35	25	-03	-17	08	-16	05	33	07	-08	-06	-06	-23	-20	-32	18	09	03	44	
36	11	17	00	48	-05	12	01	-02	-02	-09	11	-18	-12	13	11	-02	-04	45	
37	-02	23	24	-12	-25	-13	-06	06	-10	03	-05	-04	-06	20	-14	03	09	27	
38	03	06	-20	11	11	-19	-11	01	-06	41	07	-07	-01	03	-00	-12	-06	45	
39	-26	-01	15	06	40	-16	-15	08	-07	13	-11	02	03	16	01	25	09	35	
40	-11	33	-05	08	14	-01	-26	22	-05	26	-05	-01	08	-02	03	13	51	49	
41	00	09	22	25	07	17	-05	-21	05	03	15	-23	-08	05	-36	-20	-12	43	
42	-09	20	-22	-12	02	26	-31	-04	01	01	-07	00	11	-21	-20	10	-03	37	
43	-10	-61	08	04	-00	-46	03	-04	-32	00	42	-01	07	15	03	-12	08	83	
44	-12	09	16	20	-01	14	-02	16	14	-13	-05	-18	11	-22	01	17	-09	35	.52
45	-12	00	54	-05	32	-19	-18	-04	03	-05	08	-05	-18	28	-00	-28	-02	63	
46	-64	-12	-22	-20	-15	-12	05	11	-11	01	04	01	-02	01	-00	-04	02	80	
47	01	-06	07	27	09	-14	-04	13	-02	01	39	05	01	13	08	-10	19	30	
48	09	19	-01	05	-06	12	16	-25	-07	03	01	-18	02	-20	12	16	-21	24	
49	02	-04	-20	-04	42	33	01	07	07	-01	09	-08	08	-11	35	-23	05	58	
50	-03	-04	-25	-01	-27	03	-11	03	-06	08	-06	-28	-01	10	25	-11	03	42	
51	19	07	-02	-13	08	-17	-04	-15	-20	-17	-12	13	-12	07	26	-01	08	37	
52	10	-09	00	12	09	-00	22	22	-16	00	-13	13	-10	04	52	-07	-00	51	
53	-20	21	-05	57	20	-03	-10	07	-01	-07	-00	02	-00	06	03	-03	25	52	
54	-09	-11	04	18	25	14	10	51	03	-01	18	40	49	12	10	-02	01	80	
55	05	08	-02	05	-08	-04	06	07	-07	23	06	-07	-20	-06	08	10	-01	19	
56	05	13	-06	-21	02	-06	14	-01	-06	-03	-19	00	41	14	-16	-18	09	49	.71
57	08	09	03	-04	-14	24	02	53	07	08	-23	05	-18	07	-03	-01	-04	60	
58	10	08	-01	-09	-01	35	14	40	09	44	-15	03	-10	05	-04	01	-01	58	
59	08	18	01	-17	-01	-13	-07	38	-13	-23	03	01	-04	12	-12	-02	-03	50	
60	06	03	-11	-04	05	-07	-04	-18	-51	03	10	01	-03	-13	02	-12	-03	40	.73

TABLE 2, Continued

Matrix No. of Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	h^2	Reliabilities
61	06	-41	17	-46	20	-14	-05	02	12	-05	08	13	-03	-04	18	-06	-22	65	
62	10	06	-40	08	-01	44	02	02	-09	06	-10	-11	-15	-33	-02	-13	-01	45	
63	-09	05	-33	02	01	17	-10	-17	03	02	03	-16	06	-45	-16	06	07	39	
64	-06	03	-07	34	07	14	-11	02	00	-16	10	-46	05	-21	-00	-06	-04	43	.44
65	-15	-06	03	-26	-03	-12	02	-05	-02	-01	-29	16	-08	22	02	-21	00	34	
66	09	10	16	02	37	04	-01	-10	-11	-03	44	06	49	-22	-04	-06	05	77	
67	-25	14	-24	10	11	-03	-11	-03	-19	10	-07	-24	09	-00	-14	21	11	33	
68	-27	-05	-17	15	24	-00	-01	-20	12	-02	-08	-09	-09	03	14	-05	05	30	
69	09	02	22	19	03	07	-06	13	-03	-18	27	22	-23	12	03	12	27	41	
70	02	09	04	-34	-22	07	-01	07	-18	-00	03	31	22	-01	02	-07	06	45	.82
71	-03	-01	04	25	-02	12	06	-11	-08	-28	-06	-11	-10	-10	-01	-03	-03	21	
72	-04	34	07	11	-06	-00	-20	-12	-09	05	02	-04	03	-16	-11	33	-06	37	
73	-07	-04	-04	01	-04	25	-06	-02	96	-01	-07	03	04	-07	08	05	-04	105	.31
74	11	03	07	02	-30	08	21	-01	-18	-06	01	-04	-35	07	14	-12	-01	31	
75	00	-04	05	05	-02	01	05	-06	52	-01	06	02	-06	02	-03	07	-02	38	.11
76	-30	-33	00	-31	07	-16	02	-13	02	-10	-09	-24	-23	26	-02	-16	-07	54	
77	-22	19	07	21	25	11	11	-05	-12	-02	-02	-08	06	-06	14	06	07	34	
78	02	-09	-12	07	-32	30	-05	-07	12	08	-09	04	-19	-07	-05	05	47	62	.39
79	04	-10	-16	-06	-11	09	03	-15	01	-12	-12	-16	-08	-01	-08	20	28	44	.17
80	02	-06	-04	-06	57	-10	-04	-27	06	-16	15	08	12	24	02	03	-05	64	.18
81	-11	04	82	-01	-04	10	-06	26	19	-00	33	15	-21	16	03	10	08	89	.80
82	-13	-02	14	-04	10	03	-12	38	-09	-13	30	03	03	-00	13	25	14	48	.65
83	04	01	16	10	-09	23	22	06	14	-00	09	-05	-15	-11	11	-17	-02	23	.32
84	-20	56	04	06	03	-06	-20	05	-04	19	08	05	26	00	-07	06	-00	51	
85	62	06	21	18	04	08	04	-04	07	02	04	02	00	-05	-02	-10	-09	69	
86	-04	28	15	21	29	12	-23	-09	08	22	-03	-06	13	-04	13	11	55	66	
87	13	-01	18	12	-05	22	25	-09	-27	17	-05	17	04	12	-14	06	52	.71.	
88	-10	-17	-48	06	03	-00	06	11	-23	-15	-03	21	-09	00	12	-14	07	53	
89	-09	10	13	03	02	-04	-21	60	12	14	-17	33	-04	12	01	06	10	49	
90	03	02	-13	03	09	18	25	14	08	-08	-06	09	06	-08	-07	-05	-02	24	
91	-13	05	-01	-03	16	06	-05	07	14	-04	-05	09	04	02	16	-07	51	48	
92	07	02	14	-03	55	02	13	07	06	-03	-04	-12	10	05	-06	34	-13	55	
93	-10	02	17	-05	-14	-32	12	05	-29	-40	22	24	-03	26	16	-14	07	50	
94	-03	-02	07	08	-10	-02	-29	-04	-08	-18	10	09	09	-08	-09	23	-11	39	
95	-17	13	06	17	-03	14	-08	10	23	23	-05	-13	-21	-03	05	12	09	24	
96	-31	05	-06	-21	-25	-21	02	03	-11	-04	-34	-04	-19	08	-02	06	14	51	
97	03	-02	-11	03	-13	08	-01	-33	06	03	54	10	-04	-13	-03	09	02	70	
98	11	-16	01	12	-05	-21	31	-05	-07	09	06	02	-05	12	-12	-05	-01	27	
99	-02	10	-29	15	20	12	-16	-03	-08	-00	-07	25	09	-21	07	17	10	33	
100	05	-09	08	-09	04	12	06	-06	-02	69	08	07	-05	03	07	01	16	65	
101	09	-08	-02	03	07	-04	11	04	-12	49	09	-01	-02	03	05	-06	04	35	
102	08	04	17	15	-09	02	09	-18	-17	-10	01	02	08	-01	-00	08	-14	23	
103	-10	-07	05	06	-06	04	05	10	-21	-09	-14	20	08	17	08	-05	-15	32	
104	-04	-00	04	-38	15	00	-00	-07	-07	-14	-03	-03	-22	-02	-08	-06	-07	33	
105	11	-05	08	-10	-03	-07	10	05	-03	08	-00	00	-04	33	06	-01	-06	22	
106	-28	-10	-12	13	45	18	-41	03	02	26	-33	-03	-15	-07	-13	04	-00	55	
107	-25	-18	-00	16	-00	22	03	07	04	00	02	07	17	-04	14	-08	-00	35	
108	06	-28	05	-01	08	-14	05	-10	-05	-28	06	22	-23	08	-11	-29	12	39	
109	04	-02	-12	04	-12	14	26	-10	03	-02	-11	-02	05	-16	26	34	05	34	
110	-18	-50	-09	01	20	-13	-05	-02	-10	-02	11	-11	-18	25	-06	-18	06	46	
111	-08	-26	04	-03	20	-12	03	05	-13	-09	07	31	10	14	23	-03	-02	31	
112	-05	-56	-05	08	-15	-09	04	08	-11	-07	26	-11	15	02	-13	-01	-10	63	
113	01	-30	03	02	-15	06	12	24	15	13	-24	21	-07	-04	04	16	-02	35	
114	-02	11	12	03	-06	47	14	-10	-01	06	05	-17	-44	04	-13	-06	-01	63	
115	16	26	03	10	-15	-01	05	-03	20	-24	13	10	09	01	-00	-02	08	30	
116	04	-12	-02	09	13	-03	-11	-21	-15	-24	-24	-09	08	-12	-05	-09	01	31	
117	15	-01	-01	-19	01	19	20	-01	04	-15	-05	-09	06	-08	-03	-07	-16	18	
No. in + 10 - Hyperplane	69	70	63	70	66	54	68	75	73	76	69	71	68	61	66	66	88		

cases, a correlation of 0.18 is significant at the 5 per cent level and 0.24 at the 1 per cent level.) Variables 17, 19, 25, 48, 51, 55, 68, 83, 95, and 115 are not listed for any factor.

c. Below a dotted line, marker variables falling outside the ± .10 hyperplane that are not already listed and that are needed to compare with a sign pattern of loadings for the markers in the earlier studies. The matches

TABLE 3

Cosines of Angles Among Reference Vectors

Factor	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1																	
2	-08																
3	01	15															
4	-02	06	-06														
5	-18	-05	-01	00													
6	11	17	-10	15	07												
7	41	-09	05	-11	-15	09											
8	-31	05	10	-02	02	-05	-17										
9	-18	02	22	-08	01	25	-10	09									
10	-12	04	-06	-06	12	10	-15	19	04								
11	06	-07	26	06	-04	-10	03	-02	-05	-08							
12	-01	-09	12	-09	-08	-08	-03	35	04	06	-05						
13	09	09	-16	-08	13	-07	03	05	-06	-07	05	13					
14	-24	-09	41	-03	07	-52	-18	20	01	09	19	15	-13				
15	09	-12	-20	14	-05	10	10	-05	-02	-10	06	-02	-10	-04			
16	01	18	-04	-04	11	07	-04	09	04	12	-20	12	15	-27	-03		
17	-03	10	00	26	-04	-06	-34	08	-03	07	09	16	-02	15	09	07	

obtaining between study $F_1(3)$ and study $F_2(8)$ have already been discussed in (8). In any event, the loadings obtained by the marker variables in a factor from an earlier study are given at the right; if two earlier factors are shown, they have already been matched in (8).

The basic principles of factor matching have been discussed elsewhere (9). A particular problem of the present situation lies in the modification of a few of the marker variables, which are supposed to have improved intrinsic reliability but are also shortened. In general, perhaps partly through shortening and partly through regression of high loadings with the new population, the loadings of these variables are smaller here than in the earlier studies. Of course, it is expected that their loadings be lower than for the new tests, since the markers were simply to "hold the pattern" while trying out the new tests that were based on an existing perception of the true factor.

At least three markers were generally inserted for each factor to be carried into this study, two highly loaded ones and at least one other in the hyperplane. Some of these markers were used in only one of the two earlier studies. The same absolute level of loadings is not expected, and would not necessarily prove factor identity anyway. The judgments to be made, therefore, rest upon a *similarity of sign pattern over all markers* having loadings outside the hyperplane of either factor matched. (For example, with 6 markers this should occur by chance only 1 time in 32.) A secondary consideration is given to the general variance and "psychological meaning" of the factor.

Factor 1 (Table 4)

The three significantly loaded markers from the old factor—14, 44, and 46—reappear together here. Their signs are consistent, as are the signs of three other markers clear of the hyperplane in both studies—7, 19, and 22. However, $F_2$9 also has marked resemblance to the present *F*6 and must be considered an alternative there. The low fluctuation of attitudes may be an artifact here, since initial middle-of-the-road opinions leave less room for change. Therefore, it seems best to center any interpretation upon the indecisiveness and absence of bold viewpoints, and these also fit the $F_2$9 picture of cautious insecurity.

Factor 2 (Table 5)

The best over-all previous match is with $F_1$3 and $F_2$1, but the present factor is also the only one here with the $F_1$9 pattern of long reaction time, low oscillation, some loading on variable 10 and other more faintly loaded markers with the correct signs. The pattern is suggestive of the known associations of surgency, particularly the inaccuracy, low aspiration, and long reaction time (**6, 12**). If the $F_2$9 match is accepted the total picture would now also include anteversion, body sway suggestibility, suggestibility to authority and indifference to shock and disapproval (despite high awareness of errors), all psychologically consistent with central "happy-go-lucky" response of surgency (**7**).

Factor 3 (Table 6)

This factor is matched with the factors previously interpreted as the second-order factor of general ability. It has the highest loadings in the intelligence test and riddles. The failure of the logical consistency test in $F_2$2 need not concern us, for it was then new and poor. However, there are three interesting anomalies:

(1) The test of social good taste has a higher loading than the verbal-numerical test; perhaps it amounts to a test of general information that has a more appropriate difficulty level than variable 15 for the student group.

TABLE 4

Variables having Highest Loadings on Factor 1

Matrix Number	Description	Loadings Present	F_29
28	Low fluctuation of attitudes (i)	-67	
46	Little extremity of viewpoint	-64	-24
85	Indecisiveness of viewpoint	62	
31	Low fluctuation of attitudes (iv)	-59	
30	Low fluctuation of attitudes (iii)	-58	
29	Low fluctuation of attitudes (ii)	-53	
76	Breakdown of reality principle	(-30)	
107	Low scatter of self-referent time estimates	(-25)	
67	Little unreflective acceptance of unqualified statements	(-25)	
77	Overestimation of anticipated performance	(-22)	
. . .	. . . . .	. . .	. . .
14	Slow tempo	-18	-36
44	Low ratio emot. int./emot. dry reading speed	-12	-31

TABLE 5

Variables having Highest Loadings on Factor 2

Matrix Number	Description	Loadings Present	F_19
43	Little impairment on maze by shock	-61	
112	Little impairment of classification performance by disapproval	-56	
84	High inaccuracy relative to speed	56	
110	Little impairment closure performance by disapproval	-50	
9	Slow reaction time (irregular warning)	47	24
3	High rigidity in hidden words	-43	
61	Low aspiration level in coding	-41	
76	Breakdown of reality principle	(-33)	
113	Low ratio normal/shock performance on cancellation	(-30)	
. . .	. . . . .	. . .	. . .
8	Slow reaction time (regular warning)	34	24
16	Low oscillation	-30	-15
10	Low ratio reg./irreg. reaction time	-21	-18

TABLE 6

Variables having Highest Loadings on Factor 3

Matrix Number	Description	Loadings Present	F_16	F_22
81	High social good taste	82	32	
45	High logical consistency	54		-02
88	Small tendency to agree	-48		
24	High ability to solve riddles ("cope with surprises")	47	27	
15	High verbal and numerical ability	40	42	84
62	Low tendency to check remarks as "obvious"	-40		
. . .	. . . . .	. . .	. . .	. . .
114	Alertness, immediate memory	12	18	48

TABLE 7

Variables having Highest Loadings on Factor 4

Matrix Number	Description	Loadings Present	F_18	F_27
53	Decrease of myokinesis under threat	57		
33	Maturity of opinion	-56	-25	-42
36	High ratio conson./disson. recall	48	-05	-06
1	Rigidity in picture perception (I)	-47		
61	Low aspiration level in coding	-46		
2	Rigidity in picture perception (II)	-40		
104	Few jokes considered funny	-38		
64	Little considered possible for self in given time	34		
70	Little self-confidence in untried performance	-34		
. . .	. . . . .	. . .	. . .	. . .
22	High ratio emot./non-emot. recall	15	02	34
34	Suggestibility to authority	07	23	46

(2) In variables 81, 88, and 62 (which appear in Table 6), and in variables 18 and 67, this factor resembles that *K* factor which has previously been in rivalry with the *B* factor of intelligence in factorial studies of ratings (6). By comparison, $F_1 6$ and especially $F_2 2$ are free of these "taste" products of education.

(3) Factor 6 here, which was difficult to separate in rotation from the present factor, loads some of the psychomotor efficiency variables found in $F_2 2$.

Further experimental rotation may be helpful, but Factor 3 appears to be general ability. Anomaly (3) above, which will become more clear after seeing Factor 6, is probably best interpreted as a result of poor rotation of $F_2 2$.

Factor 4 (Table 7)

This is not so good a match as Factors 1 through 3, as variable 36 behaves strangely. The hyperplane is relatively poor, and perhaps an improved rotation could be found. However, the factor hangs together psychologically even in the unlisted loadings, and pictures a withdrawn, distrustful, habit-bound personality—with inner emotionality shown particularly in the $F_1 8$ and $F_2 7$ factor patterns.

Factor 5 (Table 8)

This match is still less satisfactory than that for Factor 4. The pattern of markers is consistent with $F_1 1$, $F_2 3$, or $F_1 7$, but only the latter is unmatched elsewhere in the present results and only the latter agrees with the present factor in having the highest loading of variable 39. If $F_1 7$ is considered to be rating Factor I, as speculated in (6), this picture of hostility with capability and composure would fit its obverse, as required. However, it must be emphasized that this match is a very tentative one.

Factor 6 (Table 9)

Despite having an appreciable mean variance, Factor 6 has no very high loadings. The matching of the markers is not very clear and presents two alternative sets: either with $F_1 9$ and $F_2 9$ or with $F_1 4$ and $F_2 5$. The former, correct at least as to signs, has been set out here pending further studies, despite the necessity for considering the *F*9's as possible matches also for F_1 and F_2 here.

However, the psychological meaning of Factor 6 does resemble that of the earlier factors, by combining high ideo-motor speed, reliable production of quick, complex judgments easily confused through perseveration, high rate of cube fluctuation and self-confidence. The marker for Thurstone's

TABLE 8

Variables having Highest Loadings on Factor 5

Matrix Number	Description	Loadings Present	F_17
80	High hostility of judgment	57	
92	High accuracy in Gestalt closure	55	
106	High scatter in estimates of time for others	45	
49	Slow to imitate animal sounds	42	
39	Good two-hand coordination	40	34
27	High criticism of others relative to self	-38	
50	Effect of restraint on perception	(-27)	
. . .	. . . . .	. . .	. . .
10	High ratio reg./irreg. reaction time	23	00
40	High body sway suggestibility	14	08

TABLE 9

Variables having Highest Loadings on Factor 6

Matrix Number	Description	Loadings Present	F_19	F_29
114	High efficiency of immediate memory	47		
43	Little impairment of maze performance by shock	-46	-02	-13
62	Marked tendency to check remarks as "obvious"	44		
6	Low motor rigidity (numbers)	-38		
14	Slow tempo (ideational)	-43	-06	-35
58	High number correct CMS decisions	35		
15	High verbal and numerical ability	31		
23	High ratio emot. int. to emot. dry recall	21	03	34
. . .	. . . . .	. . .	. . .	. . .
16	Low oscillation	-17	-04	-15

TABLE 10

Variables having Highest Loadings on Factor 7

Matrix Number	Description	Loadings Present	F_110
4	Low classical motor rigidity I	-65	-15
5	Low classical motor rigidity II	-61	-15
6	Low classical motor rigidity III	-48	-15
106	Small scatter in other-referent time estimates	-41	
35	Ability to state logical assumptions	33	
98	Slower decisions "on-principle" relative to decisions "in-particular"	(31)	
42	Smaller PGR for mental than for physical stimuli	(-31)	
7	High speed of judgment	(30)	
94	Little excess of aspiration over attainment	(-29)	
90	High ratio males/females liked	(25)	
. . .	. . . . .	. . .	. . .
40	Low body sway suggestibility	-26	-20

TABLE 11

Variables having Highest Loadings on Factor 8

Matrix Number	Description	Loadings Present	F_11	F_23
89	High emphasis on personal vs. institutional values	60		
57	High percent correct CMS decisions	53		
54	Big reduction of reading speed by frustration	51		
58	High number correct CMS decisions	40		
82	Good sense of poetic aptness	38		
. . .	. . . . .	. . .	. . .	. . .
18	Low questionable reading preferences	31	26	58
14	Fast tempo	31	52	55
34	Low suggestibility to authority	-26	-16	-49
16	Low oscillation	-09	-26	-34

perceptual factor *D* (**13**) was present in factor $F_1 4$, and the factor was called "psychomotor efficiency." Most of the variables loaded by the present factor could be rationalized in terms of a perceptual efficiency, with personality associations being secondary.

Factor 7 (*Table 10*)

This factor and its matching are well-defined. This takes into account that the rigidity measure of $F_1 10$ was of low reliability and could be raised to the present value by correcting for attenuation. In each study the factors matched here have the highest rigidity loading and are the only ones loading both rigidity and sway suggestibility. The latter combination was found by Eysenck (**12**) as the obverse of a factor he called "neuroticism." The present loadings of variables 35, 52, and 86 are particularly in accord with Eysenck's results. It is also noteworthy that variable 52, high frequency of flicker fusion, has come into the same factor with the three independent measures of rigidity, as required by previous theory and findings (**10**).

Factor 8 (*Table 11*)

Although the previous match of $-F_1 1$ and $-F_2 3$ is a clear one, there is some question here as to whether they match Factor 5 or 8. The latter is judged slightly better, but the former should be kept in mind in further research. Factor 5 would be good because of its high loadings on variables 10 and 33, and low loading on variable 41. However, Factor 8 is better in view of its loadings on the two highest $-F_1 1$ and $-F_2 3$ markers, variables 14 and 18, and lack of loading on variable 39. Factors 5 and 8 have opposite loadings on variable 80, hostility of judgment. Both factors load accuracy in performances requiring nice judgment, but Factor 8 shows less inhibition of expression. Among the lower loadings are high ratio of verbal to numerical ability, low agreement with obvious remarks, few jokes considered funny, and high social good taste—all consistent with the above. The whole picture suggests rating factors *K* or *A*−, the latter having previously been hypothesized for $-F_1 1$ and $-F_2 3$.

Factor 9 (*Table 12*)

This factor is essentially a doublet among the new tests; two of the three high loadings come from the same performance, and there are no loadings on markers. The sense is clearly one of depression or guilt, and may represent a mood level. The low loadings for variables 93 and 115 are also consistent with this picture.

Factor 10 (*Table 13*)

Although this factor has one of the best hyperplanes and good loadings on four markers, the pattern can be previously found only in vaguer form as

TABLE 12

Variables having Highest Loadings on Factor 9

Matrix Number	Description	Loadings Present
73	Excess adverse over favorable self-reference of events	96
75	Ratio adverse to favorable self-reference of events	52
60	Low optimism over doing good	-51
43	Low impairment of maze performance by shock	-32
32	Small myokinetic movements	(-26)
33	Immaturity of opinion	24
103	Small lag in flicker fusion	(-21)
102	Small range in flicker fusion	(-17)

TABLE 13

Variables having Highest Loadings on Factor 10

Matrix Number	Description	Loadings Present
100	High speed on CMS	69
101	Many CMS slanting lines cut	49
58	High number correct CMS decisions	44
38	High color to form in sorting	41
93	Low estimated breadth of experience	-40
22	Low emotional to non-emot. recall	-33
16	High oscillation	31
71	Low pleasant to unpleasant assns.	(-28)
20	Low initial to final performance	(-26)

TABLE 14

Variables having Highest Loadings on Factor 11

Matrix Number		Loadings Present	F_{12}	F_{24}
97	High agreement with "obvious" remarks	54		
66	High perception of threatening objects	44		
43	High maze-impairment by shock	42		
47	High fidgetometer frequency	39		09
96	Low extremity of viewpoints	-34		
81	High social good taste	33		
65	Low estimate personal prowess	-29		
. . .	. . . . .	. . .	. . .	. . .
18	Low questionable reading preferences	23	47	65
12	Low speed of closure	-20	-15	
10	High regular to irreg. reaction time	18	07	46
41	High mean PGR deflection	15	32	34

TABLE 15

Variables having Highest Loadings on Factor 12

Matrix Number	Description	Loadings Present
2	Low rigidity in picture perception (2)	57
64	Much considered possible for self in given time	-46
54	Big reduction in reading speed by frustration	40
1	Low rigidity in picture perception (1)	36
89	High personal to institutional values	33
111	High normal to shock word closure performance	(31)

TABLE 16

Variables having Highest Loadings on Factor 13

Matrix Number	Description	Loadings Present
66	High perception of threatening objects	49
54	Big reduction reading speed by frustration	49
114	Poor immediate memory	-44
56	High increase time estimate when working	41
74	Low self-reference seen in public events	-35
16	Low oscillation	-31

-5 in F_1 , and -10 in F_2 . The first three loading do depend on the same situation, but the lower loadings are the highest for their respective variables. The general picture is of impetuous performance, with some disregard of rules, oscillation, and color dominance.

Factor 11 (Table 14)

The matching with five markers is perfect as to sign, although the earlier loadings are higher. (The fidgetometer measurement is not strictly comparable.) Myokinesis, CMS circles, criticalness, and verbal-numerical ratio had slight loadings in one previous factor. The new variables in the present factor agree with the previous interpretation of self-control and general inhibition, and add a suggestion of timidity.

Factor 12 (Table 15)

This factor resembles the reverse of $F_2$10 but not $F_1$5, which was previously matched with $F_2$10. The picture of low-rigidity and high-energy would fit. This factor also has a slight resemblance to $F_1$11, but Factor 13 is a better match for it.

Factor 13 (Table 16)

Eight markers show the correct signs to match this factor with the reverse of $F_1$11, but the match is not spelled out because $F_1$11 had such small variance. The picture here, as in $F_1$11, is of poor memory, speed of judgment, and purposefulness.

Factor 14 (Table 17)

There was some difficulty in obtaining simple structure among the first two factors and this one, and these three still have the poorest hyperplanes. $F_1$3 and $F_2$1 will match the present Factors 1 and 2—especially the latter—almost as well as this factor. It seems best simply to record the present pattern with the claim of $F2$ to be an alternative match to $F_1$3 and $F_2$1, and to await the outcome of a fourth analysis before attempting to solve the rotational problem. Although the meaning of the new variables here is certainly in accord with $F_1$3 and $F_2$1 as femininity or some factor correlated with sex, the fact that the present study used an all-male sample, compared with the previously mixed sexes, may account for some of the difficulties in deciding on the proper match.

Factor 15 (Table 18)

This is a good match for $F_2$6, which also loaded fluency and suggestion of classifications. In the face of all these speed scores, the signs of variables

TABLE 17

Variables having Highest Loadings on Factor 14

Matrix Number	Description	Loadings Present	F_{13}	F_{21}
26	High criticalness of judgment	57		
63	Much considered possible for others in given time	45		
105	High increase tempo under threat	33		
21	High initial to later performance	35		
23	Low emotional to non-emot. recall	-33	13	64
62	Little tendency to consider remarks "obvious"	-33		
. . .	. . . . .	. . .	. . .	. . .
35	Unable to state logical assumptions	-32	-44	-51
37	High consonant to dissonant opinion recall	20	13	64
18	Low questionable reading preferences	19	25	57

TABLE 18

Variables having Highest Loadings on Factor 15

Matrix Number	Description	Loadings Present	F_{26}
8	Short reaction time (reg. warning)	-61	-16
52	High frequency flicker fusion	52	
9	Short reaction time (irreg. warning)	-48	-16
24	Poor performance on surprises (riddles)	-38	
41	Low PGR deflections	-36	-23
49	Slow to imitate animal sounds	35	
13	Fast alternating cube-perspective	20	31
7	Fast judgment	20	27
14	Fast tempo	-19	-34

TABLE 19

Variables having Highest Loadings on Factor 16

Matrix Number	Description	Loadings Present	F_{28}
12	Fast closure	58	
2	Low rigidity in picture perception (2)	35	
109	High normal to shock cancellation	34	
45	Low logical attitude consistency	-28	-30
39	Good two-hand coordination	25	06
41	Low PGR deflections	-20	-11
13	Fast alternating cube perspective	16	31

TABLE 20

Variables having Highest Loadings on Factor 17

Matrix Number	Description	Loadings Present
86	High static ataxia	55
91	High aspiration adaptability	51
40	High body sway suggestibility	51
78	High ego weakness (1)	47
79	High ego weakness (2)	(28)

24 and 49 are surprising. Both Factor 11 and this one resemble the "ideational rigidity" factor found as the second factor in a study of rigidity measures (10, cf. variables 1-6, 24, and 47).

Factor 16 (Table 19)

This is a fair match with $F_2$8, and only variable 39 is in one of the hyperplanes, so this is tentatively set out. The sense of the higher loadings is simply good visual perception; note that variable 2 is for the a priori key, requiring valid perceptions. However, the variance of this factor is small, and longer tests are needed for it.

Factor 17 (Table 20)

There is no match anywhere in the previous analyses for the pattern of four or five marker loadings too low to be shown in Table 20. This factor is interesting in its accounting for more variance in variables 40 and 86 than either factor 4 or 10, which might also be assumed to be "neuroticism." The whole factor portrays a timid suggestibility.

Summary

1. A total of 115 objective personality measures were factored by the multiple group method, yielding 17 factors. These were rotated to a good simple structure, roughly orthogonal.

2. The newly-designed tests, based upon hypotheses about factors from previous analyses, generally have higher loadings than the older tests used as "marker variables." However, the latter have enough communality to "hold the pattern" and permit matching new with old factors.

3. Thirteen of the present factors are matched with earlier factors, nine confidently and four tentatively. This leaves only two of the F_1 series and one of the F_2 series unmatched. Since $F_1$5 and $F_2$10 are themselves a match, the tests used in the three series appear to involve nineteen factors.

4. The psychological consistency of each factor is good, but space has not permitted a detailed discussion here.

5. It is now possible to undertake the intensive study of individual factors (without using the large batteries so far necessary to avoid confusing similar-appearing factors) to find more valid, highly-loaded measures. The best factors for this purpose, defined and confirmed in the present study, are factors 1 ($F_2$9), 2 ($F_1$8), $F_2$7), 3 ($F_1$6, $F_2$2), 7 ($F_1$10), 11 ($F_1$2, $F_2$4), 15 ($F_2$6), 16 ($F_2$8), and possibly 6 ($F_1$4, $F_2$5).

REFERENCES

1. Cattell, R. B. The riddle of perseveration, I and II. *J. Person.* 1946, **14**, 229-267.
2. Cattell, R. B. The description and measurement of personality. New York: World Book, 1946.

3. Cattell, R. B. Primary personality factors in the realm of objective tests. *J. Person.*, 1948, **16**, 459-487.
4. Cattell, R. B. A guide to mental testing. 3rd edit. Inst. Pers. and Ability Testing, 1608 Coronado Dr., Champaign, Ill., 1954.
5. Cattell, R. B. The primary personality factors in women compared with those in men. *Brit. J. Psychol., Stat. Sect.*, 1948, **1**, 114-133.
6. Cattell, R. B. Personality: a systematic theoretical and factual study. New York: McGraw-Hill, 1950.
7. Cattell, R. B. The main personality factors in questionnaire, self estimate material. *J. soc. Psychol.*, 1950, **31**, 3-38.
8. Cattell, R. B. A factorization of tests of personality source traits. *Brit. J. Psychol. Stat. Sect.*, 1951, **4**, 165-178.
9. Cattell, R. B. Factor analysis for psychologist and social scientist. New York: Harper Bros, 1952.
10. Cattell, R. B. and Tiner, L. G. The varieties of structural rigidity. *J. Person.* 1949, **17**, 321-341.
11. Cattell, R. B., Dubin, S. S., and Saunders, D. R. Personality structure in psychotics by factorization of clinical tests. *J. ment. Sci.*, 1954, **100**, 154-176.
12. Eysenck, H. J. Dimensions of personality. London: Kegan Paul, 1947.
13. Thurstone, L. L. A factorial study of perception. Chicago: Univ. of Chicago Press, 1944.
14. Thurstone, L. L. Multiple-factor analysis. Chicago: Univ. of Chicago Press, 1949.

The writers wish to express their indebtedness for support of this project to the National Institute of Mental Health, of the National Institutes of Health, Public Health Service, Washington, D.C.

Manuscript received 6/8/52
Revised manuscript received 10/8/53

PSYCHOMETRIKA—VOL. 19, NO. 3
SEPTEMBER, 1954

NOTE ON CONSTRUCTION OF AN ITEM ANALYSIS TABLE FOR THE HIGH-LOW-27-PER-CENT GROUP METHOD

C. T. FAN

EDUCATIONAL TESTING SERVICE

This paper describes the construction of a new item analysis table for the high-low-27-per-cent group method. The table provides a ready means of translating the observed proportions of success in the two extreme groups (p_H , p_L) into measures of item difficulty and item discrimination (p, Δ, and r). The tabled values of both the difficulty index, p, and the discrimination index, r, have been derived from Karl Pearson's tables of the normal bivariate surface.

The present paper describes the construction of a new item analysis table for the high-low-27-per-cent group method. (The table has been published by Educational Testing Service, Princeton, New Jersey. The price per copy is one dollar, which must accompany the order, to save the cost and trouble of accounting and billing.) The basic data are two measures for each test item, namely, the proportion, p_H , of correct responses among the 27 per cent of the group scoring highest on the criterion measure and the proportion, p_L , of correct responses among the 27 per cent of the group scoring lowest on the criterion measure. Entering the table with p_H and p_L , one finds three statistics, as follows: (1) item-difficulty index, p, which is the estimated proportion of success in the total sample; (2) item-discrimination index, r, which is the estimated correlation between the criterion score and the continuous score assumed to underlie responses to the item; (3) a second item-difficulty index, Δ, which is determined from p and expressed in terms of normal curve deviate with mean of 13 and standard deviation of 4.

The tabled values of both the difficulty index, p, and the discrimination index, r, have been derived from Karl Pearson's tables of the normal bivariate surface. It may be noted that the values of p thus obtained differ systematically from those estimated by averaging the values of p_H and p_L , as has been done in other tables of the present type. The mean of p_H and p_L overestimates values of p less than .50 and underestimates values of p greater than .50, particularly in the case of extreme values of p and high values of r.

The procedure involved in setting up the tables included two major steps: The first step was to determine the values of p_H and p_L for the specified values of r and p provided by Pearson's tables. The second step was to find the values of p and r for given values of p_H and p_L . These two steps will be described next.

I. *Determining the Values of p_H and p_L for Specified Values of r and p*

In Table VIII of Pearson's *Tables for Statisticians and Biometricians* (*Part II*) volumes of the normal bivariate surface, denoted as d/N, are listed corresponding to the normal deviates from .0 to 2.6, in intervals of .1, for twenty-one specified r's (from .00 to + 1.00, in intervals of .05). Here d and N are shown in a four-fold table:

TABLE 1

c	d	$c + d$
a	b	$a + b$
$a + c$	$b + d$	N

According to the definition given for Pearson's tables this four-fold table should be so arranged that a is the quadrant in which the centroid of the correlation surface falls; then $c + d$ and $b + d$ are each equal to or less than 50 per cent of the total sample.

For the high and low 27-per-cent groups a similar table can be arranged to show the relations of p_H , p_L , and p for a specified r:

TABLE 2

Low Group (27%)	Middle Group (46%, omitted)	High Group (27%)	Total Sample (100%)
p_L	(p_M)	p_H	p
q_L	(q_M)	q_H	q
1.00	(1.00)	1.00	1.00

When we examine these two tables, it is clear that d/N (where N is the number of cases in the total sample) in Table 1 and p_H (based on the high 27-per-cent group) in Table 2 can be made comparable if we assign a value of .27 to $(b + d)/N$ and redefine as follows: 1) $(c + d)/N$ in Table 1 defined as p in Table 2, 2) c in Table 1 defined as $p_L + p_M$ in Table 2, and 3) a in Table 1 defined as $q_L + q_M$ in Table 2. Moreover, the normal deviate values, k, in Pearson's Table VIII can be taken as the normal deviates corresponding to the values of p, and the values of d/N above the point where the normal deviate $h = .612813$ can be taken as the number of correct responses among the high 27-per-cent group divided by the number of cases in the total sample. These values of d/N can be obtained from Table VIII by interpolation and can be converted to p_H by dividing by .27.

Pearson's Table IX is the same as his Table VIII except that the listed values of d/N are for negative correlations. It is possible to change the correlation to a positive value by reversing the distribution of the criterion scores and treating the fixed point at the normal deviate, $h = .612813$, as the cutting point for the low 27-per-cent group. The values of d/N below this point can then be taken as the correct responses among the low group in proportion to the total sample. Thus values of d/N obtained from Table IX can be converted to p_L.

The corresponding values of p_H and p_L can be matched after they are computed from Table VIII and Table IX for the same specified values of p and r. For example, for $r = .55$, the values of p_H and p_L corresponding to their specified values of p are obtained as follows:

k	p	p_H	p_L	p'	p'_H	p'_L
0.0	.5000	.7777	.2223	.5000	.7777	.2223
0.1	.4602	.7420	.1896	.5398	.8104	.2580
0.2	.4207	.7037	.1600	.5793	.8400	.2963
0.3	.3821	.6629	.1336	.6179	.8664	.3371
...	.....	.....	.....	.....	.....	.....
1.6	.0548	.1459	.0044	.9452	.9956	.8541
1.7	.0446	.1218	.0031	.9554	.9969	.8782
1.8	.0359	.1006	.0022	.9641	.9978	.8994
1.9	.0287	.0823	.0015	.9713	.9985	.9177

The values of p_H and p_L (for each r) can be obtained from Pearson's tables only for $p \leq .5000$; however, the values of p'_H and p'_L for $p' > .5000$ can easily be obtained by using the relation that when $p' = 1 - p$, $p'_H = 1 - p_L$, and $p'_L = 1 - p_H$. This relation can be shown by Table 2 if both coordinates are reversed in direction when $q = p'$, $q_L = p'_H$, and $q_H = p'_L$.

It was found in the construction of the tables that linear interpolation was inadequate for determining p_H and p_L values corresponding to $h = .612813$. When a cubic interpolation was applied, however, satisfactory results were obtained. The following equation was derived:

$$y = -.034830\, y_1 + .920515\, y_2 + .135330\, y_3 - .021015\, y_4 ,$$

where y is the desired entry corresponding to $h = .612813$ and y_1, y_2, y_3, y_4 are the d/N entries corresponding to h_i of .5, .6, .7, .8, respectively. The needed interpolation coefficients may be shown to be the row vector resulting from multiplying the following two matrices:

$$[1 \quad h \quad h^2 \quad h^3] \cdot \begin{bmatrix} 1 & h_1 & h_1^2 & h_1^3 \\ 1 & h_2 & h_2^2 & h_2^3 \\ 1 & h_3 & h_3^2 & h_3^3 \\ 1 & h_4 & h_4^2 & h_4^3 \end{bmatrix}^{-1} .$$

TABLE 3

Proportions of Success in the Highest and Lowest 27 Per Cent in a Normal Bivariate Population

x	p	p_H	p_L	p_H	p_L	p_H	p_L	p_H	p_L	p_H	p_L
		r = .05		r = .10		r = .15		r = .20		r = .25	
0.0	.5000	.5244	.4756	.5489	.4511	.5735	.4265	.5981	.4019	.6229	.3771
0.1	.4602	.4845	.4359	.5090	.4117	.5336	.3874	.5584	.3632	.5835	.3388
0.2	.4207	.4448	.3969	.4690	.3731	.4934	.3494	.5181	.3258	.5431	.3022
0.3	.3821	.4056	.3588	.4293	.3358	.4533	.3129	.4776	.2901	.5023	.2674
0.4	.3446	.3673	.3222	.3903	.3000	.4136	.2781	.4374	.2563	.4615	.2348
0.5	.3085	.3303	.2871	.3524	.2661	.3748	.2453	.3977	.2247	.4211	.2045
0.6	.2743	.2948	.2540	.3158	.2342	.3373	.2147	.3592	.1955	.3816	.1767
0.7	.2420	.2613	.2230	.2810	.2045	.3013	.1864	.3220	.1687	.3432	.1514
0.8	.2119	.2298	.1943	.2482	.1772	.2671	.1606	.2865	.1443	.3064	.1286
0.9	.1841	.2006	.1680	.2176	.1524	.2350	.1372	.2529	.1225	.2714	.1083
1.0	.1587	.1737	.1441	.1892	.1299	.2052	.1163	.2216	.1031	.2386	.0904
1.1	.1357	.1492	.1226	.1632	.1099	.1777	.0977	.1926	.0860	.2080	.0748
1.2	.1151	.1272	.1034	.1397	.0922	.1527	.0814	.1661	.0711	.1799	.0614
1.3	.0968	.1075	.0865	.1186	.0767	.1301	.0673	.1420	.0583	.1543	.0499
1.4	.0808	.0901	.0718	.0999	.0632	.1100	.0551	.1205	.0474	.1313	.0402
1.5	.0668	.0749	.0591	.0834	.0517	.0922	.0447	.1013	.0382	.1107	.0321
1.6	.0548	.0618	.0482	.0690	.0419	.0766	.0360	.0845	.0305	.0926	.0254
1.7	.0446	.0505	.0390	.0567	.0337	.0631	.0287	.0698	.0241	.0768	.0199
1.8	.0359	.0409	.0312	.0461	.0268	.0516	.0227	.0572	.0189	.0631	.0154
1.9	.0287	.0328	.0248	.0372	.0212	.0417	.0178	.0465	.0147	.0514	.0119
		r = .30		r = .35		r = .40		r = .45		r = .50	
0.0	.5000	.6480	.3520	.6732	.3268	.6988	.3012	.7247	.2753	.7510	.2490
0.1	.4602	.6088	.3144	.6345	.2898	.6606	.2651	.6872	.2401	.7143	.2150
0.2	.4207	.5685	.2786	.5944	.2549	.6207	.2313	.6476	.2076	.6752	.1838
0.3	.3821	.5275	.2449	.5532	.2224	.5795	.2000	.6065	.1777	.6342	.1556
0.4	.3446	.4862	.2135	.5115	.1923	.5374	.1714	.5641	.1507	.5916	.1304
0.5	.3085	.4451	.1846	.4696	.1649	.4949	.1456	.5210	.1266	.5480	.1081
0.6	.2743	.4045	.1582	.4282	.1401	.4525	.1225	.4776	.1053	.5038	.0887
0.7	.2420	.3650	.1345	.3875	.1180	.4106	.1021	.4346	.0867	.4595	.0720
0.8	.2119	.3268	.1133	.3480	.0985	.3698	.0843	.3924	.0707	.4159	.0578
0.9	.1841	.2904	.0946	.3101	.0814	.3304	.0689	.3514	.0570	.3732	.0459
1.0	.1587	.2560	.0783	.2741	.0667	.2928	.0558	.3121	.0456	.3322	.0361
1.1	.1357	.2239	.0642	.2403	.0541	.2573	.0447	.2748	.0360	.2930	.0280
1.2	.1151	.1942	.0522	.2090	.0435	.2242	.0355	.2400	.0282	.2562	.0215
1.3	.0968	.1671	.0420	.1802	.0346	.1937	.0279	.2077	.0218	.2221	.0164
1.4	.0808	.1425	.0335	.1540	.0273	.1659	.0217	.1781	.0167	.1907	.0123
1.5	.0668	.1205	.0264	.1305	.0213	.1408	.0167	.1514	.0126	.1622	.0091
1.6	.0548	.1010	.0207	.1096	.0165	.1184	.0127	.1275	.0094	.1366	.0067
1.7	.0446	.0839	.0160	.0912	.0126	.0987	.0096	.1063	.0070	.1140	.0049
1.8	.0359	.0691	.0123	.0752	.0095	.0815	.0071	.879	.0051	.0943	.0035
1.9	.0287	.0564	.0094	.0615	.0072	.0667	.0053	.719	.0037	.0771	.0025
		r = .55		r = .60		r = .65		r = .70		r = .75	
0.0	.5000	.7777	.2223	.8049	.1951	.8326	.1674	.8609	.1391	.8897	.1103
0.1	.4602	.7420	.1896	.7705	.1640	.7998	.1382	.8301	.1122	.8614	.0862
0.2	.4207	.7037	.1600	.7330	.1363	.7635	.1126	.7952	.0892	.8285	.0663
0.3	.3821	.6629	.1336	.6927	.1119	.7238	.0905	.7566	.0698	.7913	.0501
0.4	.3446	.6202	.1103	.6500	.0908	.6813	.0719	.7144	.0539	.7499	.0371
0.5	.3085	.5760	.0901	.6054	.0727	.6364	.0563	.6693	.0409	.7047	.0271
0.6	.2743	.5310	.0727	.5595	.0576	.5896	.0434	.6217	.0306	.6563	.0193
0.7	.2420	.4855	.0581	.5128	.0450	.5417	.0331	.5724	.0225	.6055	.0136
0.8	.2119	.4404	.0458	.4661	.0347	.4932	.0248	.5221	.0163	.5532	.0093
0.9	.1841	.3960	.0357	.4199	.0265	.4450	.0184	.4717	.0116	.5002	.0063
1.0	.1587	.3531	.0275	.3749	.0199	.3978	.0134	.4219	.0081	.4476	.0042
1.1	.1357	.3119	.0209	.3316	.0148	.3521	.0096	.3736	.0056	.3962	.0027
1.2	.1151	.2731	.0157	.2906	.0108	.3087	.0068	.3274	.0038	.3468	.0017
1.3	.0968	.2369	.0117	.2521	.0078	.2678	.0048	.2839	.0025	.3003	.0011
1.4	.0808	.2035	.0086	.2167	.0056	.2301	.0033	.2436	.0016	.2571	.0007
1.5	.0668	.1732	.0062	.1843	.0039	.1956	.0022	.2068	.0011	.2177	.0004
1.6	.0548	.1459	.0044	.1553	.0027	.1645	.0015	.1736	.0007	.1823	.0002
1.7	.0446	.1218	.0031	.1295	.0019	.1370	.0010	.1442	.0004	.1509	.0001
1.8	.0359	.1006	.0022	.1068	.0013	.1129	.0006	.1185	.0003	.1236	.0001
1.9	.0287	.0823	.0015	.0873	.0008	.0920	.0004	.0963	.0002	.1001	.0000
		r = .80		r = .85		r = .90		r = .95		r = 1.00	
0.0	.5000	.9189	.0811	.9478	.0522	.9749	.0251	.9955	.0045	1.0000	.0000
0.1	.4602	.8938	.0607	.9272	.0364	.9607	.0153	.9902	.0019	1.0000	
0.2	.4207	.8637	.0445	.9011	.0247	.9407	.0090	.9805	.0007	1.0000	
0.3	.3821	.8285	.0319	.8690	.0163	.9139	.0051	.9640	.0003	1.0000	
0.4	.3446	.7883	.0224	.8309	.0104	.8795	.0027	.9382	.0001	1.0000	
0.5	.3085	.7434	.0153	.7867	.0065	.8372	.0014	.9013	.0000	1.0000	
0.6	.2743	.6943	.0103	.7370	.0039	.7873	.0007	.8522		1.0000	
0.7	.2420	.6419	.0067	.6827	.0023	.7306	.0003	.7918		.8963	
0.8	.2119	.5871	.0043	.6249	.0013	.6686	.0001	.7222		.7848	
0.9	.1841	.5311	.0027	.5650	.0007	.6031	.0001	.6469		.6819	
1.0	.1587	.4750	.0016	.5044	.0004	.5363	.0000	.5695		.5878	
1.1	.1357	.4199	.0010	.4447	.0002	.4702		.4938		.5026	
1.2	.1151	.3668	.0006	.3871	.0001	.4067		.4223		.4263	
1.3	.0968	.3168	.0003	.3328	.0000	.3472		.3570		.3585	
1.4	.0808	.2703	.0002	.2827		.2928		.2985		.2993	
1.5	.0668	.2281	.0001	.2372		.2441		.2472		.2474	
1.6	.0548	.1902	.0001	.1968		.2013		.2029		.2030	
1.7	.0446	.1568	.0000	.1614		.1642		.1650		.1652	
1.8	.0359	.1279		.1310		.1327		.1331		.1331	
1.9	.0287	.1032		.1052		.1062		.1064		.1064	

The cubic interpolation of d/N values and the conversion to p_H and p_L were performed by means of punched-card procedures which yielded a set of values of p_H and p_L corresponding to the specified values of p and r. Since this set of values can be useful to those who are interested in the normal bivariate surface and to those who may seek more accurate values of r and p (to three figures), which can be derived directly from this table, it is given with this note as Table 3. Table 3 presents values of p_H and p_L only for $.5000 \geq p \geq .0287$. Values of p_H and p_L for $.9713 \geq p \geq .5000$ can be obtained directly from the table in the manner described earlier but have been omitted for economy of space. Values for $p < .0287$ and for $p > .9713$ are not given because of their unreliability for item analysis purposes. For the same reason no values for p outside the range, .05 to .95, appear in the final table.

II. *Finding the Values of p and r for Given Values of p_H and p_L*

Given Table 3, the problem now becomes that of finding p and r for given values of p_H and p_L , which will be written to the nearest hundredth. Here, there is another problem of interpolation, which could be accomplished in a number of different ways. The procedure described below was selected because it promised to yield satisfactorily accurate results.

First, p_H was plotted against p for each value of r in Table 3. Figure 1 is presented to illustrate the graph and contains seven of the original twenty-one curves. The actual working graph was drawn on a larger scale, with the result that it was possible to read the figures to three places on both coordinates. Then values of p were read from the graph for specified values of p_H from .01 to .99 and r in intervals of .05, as given.

In like manner, p_L could be plotted against p for each r of Table 3 for reading corresponding values of p for specified values of p_L and r. The resulting plot, however, would be exactly the same graph as Figure 1 with both coordinate scales reversed in direction and with the ordinate representing p_L . Therefore, the values of p for specified values of p_L and r can be derived from the table just described through the relation of $p' = 1 - p$ when $p_L' = 1 - p_H$.

At this point the values of p were known for given values of p_H and r, as were the values of p for given values of p_L and r. The next step was to put this information together in such a way that p and r could be found for given values of p_H and p_L . This step was accomplished by means of a second graph (Figure 2). It will be recalled that twenty-one pairs of p and r values were obtained from Figure 1 for every desired value of p_H . In Figure 2 the coordinate axes represent r and p. Each p_H is represented by a curve drawn through the twenty-one points. Similar curves represent the p_L values. The coordinates of the intersection of any two of these curves are the corresponding r and p values. The working graph which was used to read the final figures was far more complete than Figure 2, which serves only to illustrate the

procedure. The working graph included curves for p_L values from .01 to .30 in steps of .01 and from .30 to .99 in steps of .02, and for p_H values from .01 to .70 in steps of .02 and from .70 to .99 in steps of .01. The curves for odd values of p_L greater than .30 and for odd values of p_H less than .70 were omitted because linear interpolation was found to be adequate in those

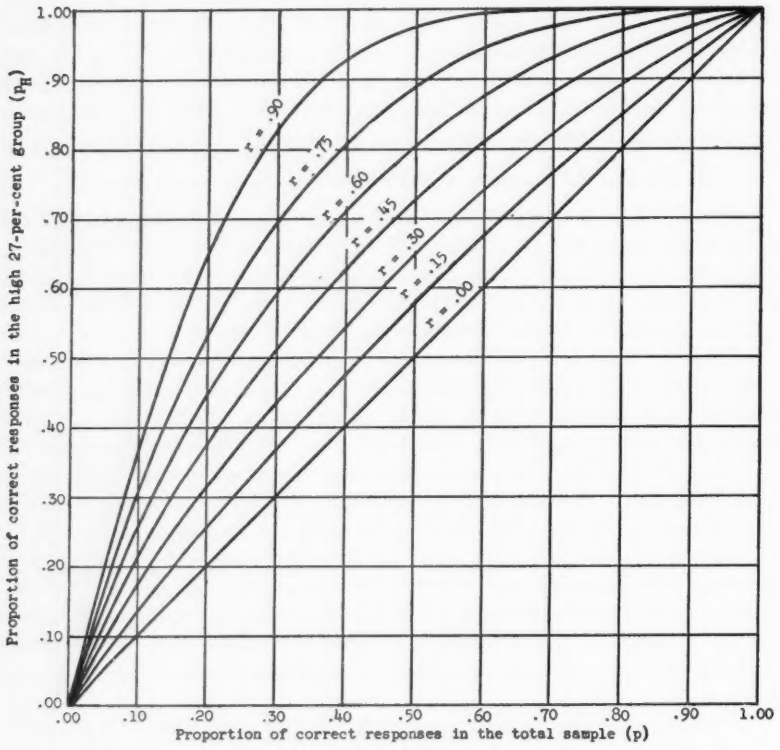


Figure 1

Simplified Version of Chart Used for Graphic Estimate of p for Given Values of p_H and r

areas of the graph. The graph was drawn to so large a scale that both p and r could be read to three places. The symmetry of the graph about the line $p = .50$ afforded a convenient check on the plotting of the curves. Unlike other graphs of this type, which require interpolation between curves on a nonlinear scale, the present graphs make it possible to read values on the

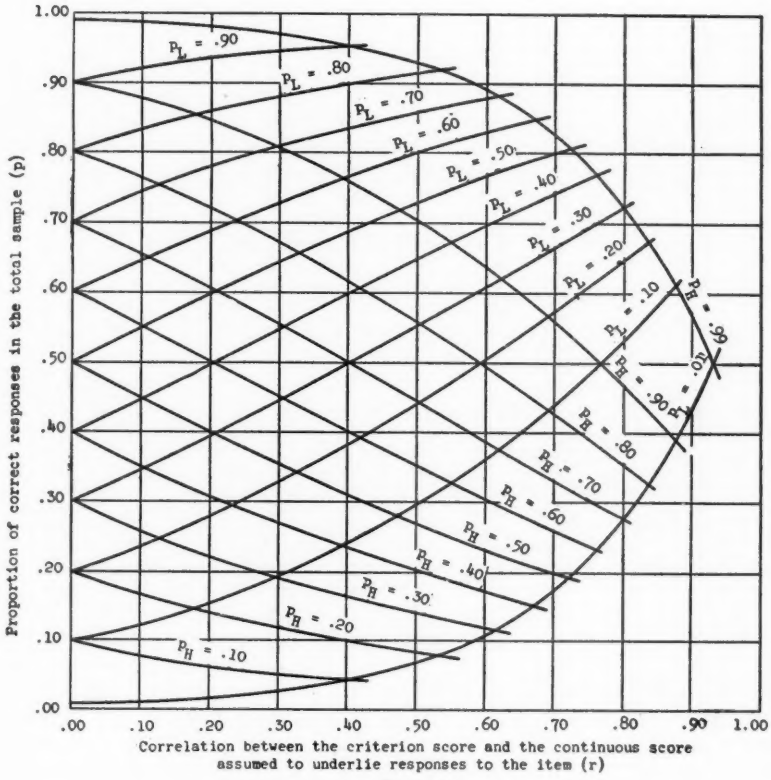


Figure 2

Simplified Version of Chart Used for Graphic Estimate of p and r for Given Values of p_H and p_L

linear scales of the coordinate axes. The linear scale is, of course, far superior to the nonlinear scale from the point of view of convenience and accuracy.

The delta values were computed from the normal deviates corresponding to the p values. Finally, all values of p and r were rounded to two figures and values of delta were kept to one decimal place.

Manuscript received 9/1/53
Revised manuscript received 11/13/53

PSYCHOMETRIKA—VOL. 19, NO. 3
SEPTEMBER, 1954

CHANGES IN FACTOR STRUCTURE OF A COMPLEX PSYCHOMOTOR TEST AS A FUNCTION OF PRACTICE

EDWIN A. FLEISHMAN AND WALTER E. HEMPEL, JR.

AF PERSONNEL AND TRAINING RESEARCH CENTER*
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

Scores obtained at eight different stages of practice on the Complex Coordination Test together with scores on 18 reference tests were subjected to a Thurstone Centroid Factor Analysis. Nine meaningful factors were identified in the experimental battery. The results indicated considerable, but systematic, changes in the factor structure of the Complex Coordination Test as practice on the task was continued. The test became less complex (factorially) as practice was continued. Moreover, there was a change in the *nature* of the factors contributing variance at early and later stages of practice. Implications of the findings are related to certain problems of learning theory, psychomotor test development, and criterion analysis.

Introduction

In some previous studies (**3**, **8**) evidence was found that even during the short time of administration of a single psychomotor test, the ability or abilities sampled may shift materially in importance. This was indicated by variations in the factor patterns and in the weights of particular factors at various stages of practice on the particular tests investigated. Similar results also had been found earlier in factorial studies of extended practice on printed tests (**2**, **5**, **13**, **15**, **16**, **17**). The problem appears more crucial in the psychomotor area, however, since measures of such skills typically show quite rapid and extensive improvement with even brief amounts of practice. It then becomes important to establish what abilities are being sampled at different stages in performance as practice is continued on particular psychomotor tasks. Such knowledge would have implications for future test development in this aptitude area as well as for questions concerning the processes involved in the learning of complex perceptual-motor skills.

*Skill Components Research Laboratory. The opinions or conclusions contained in this report are those of the authors and do not necessarily reflect the views or indorsement of the Department of the Air Force.

The writers are indebted to Dr. Jack A. Adams for the basic data on which this analysis is based.

Purpose

The study to be reported here was designed to investigate further the nature of these changes in ability patterns which seem to occur as practice continues on a complex psychomotor test. More specifically, an attempt was made to *identify* the factors involved at different stages of performance on a criterion psychomotor task, through the inclusion of well-established reference variables in the analysis. The interpretation of the nature of the factors involved at various practice stages has been limited in the earlier studies by the absence of such sufficiently well-defined reference variables. Moreover, in the present study practice on the criterion task was continued over a considerably longer time period than had been investigated previously.

The analysis also was aimed at identifying (1) the stages of practice in which the task is most complex (in terms of the number of abilities sampled), (2) the stages at which systematic changes in factor structure occur, (3) the stage at which the factor structure becomes stabilized, (4) the relative importance of "motor" versus "non-motor" factors at early and late stages of practice.

Method

A factor analysis of performance at various stages of practice on the criterion task together with selected reference tests was carried out by the Thurstone Centroid Method (**12**). The analysis is based on a sample of 197 basic airmen tested at Lackland Air Force Base. Variables in the analysis are described below.

The Criterion Practice Task: The Complex Coordination Test, Model E, was the criterion task learned. The test has been described in detail elsewhere by Melton (**8**). Essentially, the subject is required to make complex motor adjustments of an airplane-type stick and rudder in response to successively presented patterns of visual signals. A correct response (movement of stick and rudder controls to proper positions) is not accomplished until both the hands and feet have completed the appropriate movements. A new pattern appears as each correct response is completed. Score is the number of correct responses completed in a given test period. This task was selected since it has a high learning ceiling, shows no appreciable decrease in relative variability between subjects as practice is continued, and was suspected to be factorially complex. The acquisition curve for this task has been presented elsewhere (**1, 10**).

Practice on the test was continued over 64 two-minute trials. For each subject, testing was accomplished over a two-day period, each day involving a morning and afternoon session. Each session included 16 trials separated by one-minute rest intervals.

Scores obtained during the following eight segments of the total practice period were selected for inclusion in the present analysis:

1. Stage 1—Trials 1-5.
2. Stage 2—Trials 12-16.
3. Stage 3—Trials 17-21.
4. Stage 4—Trials 28-32.
5. Stage 5—Trials 33-37.
6. Stage 6—Trials 44-48.
7. Stage 7—Trials 49-53.
8. Stage 8—Trials 60-64.

These stages of practice on the criterion task, which represent the first eight variables in the present analysis, include the first and last ten minutes of practice during each of the four testing sessions.

Test Variables: The following printed tests were included in the analysis. They have been described in detail by Guilford (7).

9. *Numerical Operations—II.* A highly speeded test requiring simple subtraction and division operations.
10. *Dial and Table Reading.* A series of instrument dials and mathematical tables from which information must be read accurately and quickly.
11. *Mechanical Principles.* Pictorial items which require the comprehension of principles and mechanisms, such as leverage, and rotation and transformation of motion, involved in the action and uses of various mechanical devices.
12. *General Mechanics.* Verbally presented items of practical mechanical information dealing with the use and operation of familiar mechanical methods and devices.
13. *Speed of Identification.* Pictorial items in which the silhouette of an object must be identified quickly when it is rotated and imbedded in a group of similar silhouettes.
14. *Pattern Comprehension.* A series of drawings requiring visualization of relationships between components of solids and their unfolded flat projections.
15. *Visual Pursuit.* From a series of mazes of irregularly curved lines, the task is to trace each line visually from its beginning to its proper termination point.
16. *Decoding Test.* A series of short words, written in a code of flag symbols, must be decoded by relating the position of repeated symbols to the positions of repeated letters.
17. *Instrument Comprehension.* For each item which presents views of cockpit instruments the examinee must determine the proper position or orientation of an airplane.
18. *Spatial Orientation.* From a large aerial photograph or map, the examinee must find the area that matches each of a series of small photographs.
19. *Speed of Marking.* The examinee marks an IBM answer sheet in the indicated spaces as rapidly as possible.
20. *Log Book Accuracy.* On an IBM answer sheet, the examinee marks as quickly and accurately as possible the proper letter indicated in a separate test booklet.

The following apparatus tests were included. They have been described in detail by Melton (8).

21. *Rotary Pursuit.* The examinee attempts to keep a stylus in contact with a small metallic target set in a rapidly revolving disk.
22. *Plane Control.* The task is to make compensatory adjustments of stick and pedal controls in order to keep a model plane straight and level.
23. *Discrimination Reaction Time.* The examinee manipulates one of four toggle switches as quickly as possible in response to a series of visual stimulus patterns differing from one another with respect to the spatial arrangement of their component parts.

Three additional apparatus tests which have not been described elsewhere were also included.

24. *Nut and Bolt.* The task is to insert and fasten as many nut and bolt assemblies as possible through a series of holes in an upright metal plate.
25. *Reaction Time.* A series of reactions in which the examinee must move his hand six inches as quickly as possible to a switch in response to a light stimulus which appears at varying intervals.
26. *Rate of Movement.* The examinee's task is to break the beams between a series of photoelectric cells, one after another, by making scalloped movements of his hand as rapidly as possible.

Results

Interpretation of Factors: The intercorrelations (Pearson product-moment) among the 26 variables are presented in Table 1. Ten factors were extracted from this matrix. The complete centroid factor matrix obtained is presented in Table 2. Orthogonal rotations were made using Zimmerman's graphical method (**20**). Table 3 presents the orthogonal solution of rotated factor loadings obtained using the criteria of simple structure and positive manifold.

Variables having orthogonal projections of .30 or larger on the rotated axes were considered in defining a factor. However, loadings of .25 or higher are considered significant (see Table 3). Interpretations given certain of the factors have drawn on several previous Air Force analyses (**6**, **7**, **9**, **18**, **19**).

Factor I is common only to stages of practice on the criterion task, the Complex Coordination Test. Moreover, it can be seen (Table 3) that the loadings increase progressively through stage 4 and then remain at about the same level through the remaining practice stages. For the present, this factor is called the *Complex Coordination Test Specific.*

Factor II is common to all eight stages of practice on the Complex Coordination Test and to the Rotary Pursuit Test. It also shows a low loading in the Plane Control Test. This factor appears to be the same as the factor called *Psychomotor Coordination* in many previous analyses of the Aircrew Classification Batteries. It has consistently been identified in all analyses which have included the Rotary Pursuit and Complex Coordination Tests (the latter represented by stage 1 in the present analysis). The factor has

TABLE 1

Intercorrelations among Test and Practice Task Variables*

Test	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1. Practice Stage 1, Complex Coordination	-	75	73	66	64	57	63	59	28	51	49	38	50	55	43	50	44	47	46	33	51	40	50	36	08	25
2. Practice Stage 2, Complex Coordination		-	85	85	84	79	77	79	30	46	40	31	45	41	47	38	38	39	41	32	51	45	44	30	22	32
3. Practice Stage 3, Complex Coordination			-	85	83	79	81	79	30	45	39	29	46	44	45	39	42	46	44	35	59	44	48	31	27	31
4. Practice Stage 4, Complex Coordination				-	90	88	86	85	26	40	36	27	41	38	42	34	35	40	40	34	51	44	42	32	30	28
5. Practice Stage 5, Complex Coordination					-	90	87	86	22	37	36	25	37	37	39	32	31	38	36	33	50	42	38	28	30	34
6. Practice Stage 6, Complex Coordination						-	85	86	23	34	29	20	31	27	34	24	27	31	33	28	47	39	33	25	27	37
7. Practice Stage 7, Complex Coordination							-	90	23	36	33	25	32	33	36	29	28	39	39	31	50	39	35	24	33	30
8. Practice Stage 8, Complex Coordination								-	24	34	30	24	35	29	38	29	30	35	34	28	49	36	33	26	27	32
9. Numerical Operations - II									-	63	32	23	44	34	23	45	42	31	37	40	22	08	39	13	09	12
10. Dial and Table Reading										-	54	43	61	59	40	61	51	54	49	50	38	22	54	28	05	24
11. Mechanical Principles											-	52	46	52	34	43	30	44	33	26	27	22	37	27	-05	12
12. General Mechanics												-	29	30	17	32	30	32	17	23	26	29	30	30	02	04
13. Speed of Identification													-	57	44	58	40	58	55	48	31	23	52	31	15	22
14. Pattern Comprehension														-	43	60	44	49	45	39	32	31	54	30	03	26
15. Visual Pursuit															-	33	35	46	32	32	29	20	38	24	10	16
16. Decoding																-	53	59	41	32	24	19	55	22	08	18
17. Instrument Comprehension																	-	46	34	22	28	29	47	31	11	24
18. Spatial Orientation																		-	49	37	40	21	41	26	06	15
19. Speed of Marking																			-	55	26	18	43	20	14	28
20. Log Book Accuracy																				-	27	19	33	18	23	26
21. Rotary Pursuit																					-	35	34	28	26	28
22. Plane Control																						-	36	20	20	20
23. Discrimination Reaction Time																							-	22	23	23
24. Nut and Bolt																								-	05	14
25. Reaction Time																									-	30
26. Rate of Movement																										-

* Decimal points omitted.

been defined rather broadly as representing either coordination of the large muscles of the body, in movements of moderate scope, or coordination of such movements with the perception of a visual stimulus.

The present analysis indicates that this factor is most adequately sampled in stages 2 and 3 of the Complex Coordination Test, but that it maintains its importance at all stages of practice on the test.

Factor III is defined primarily by the two apparatus tests, Reaction Time and Rate of Movement, which are practically pure measures of this factor. Significant loadings are also evidenced by the Rotary Pursuit, Plane Control, and Discrimination Reaction Time Tests. It seems clear that this factor may be defined as *Rate of Movement.* It is possible that this factor extends beyond rate of arm movement to leg and feet movements as well.

With respect to the importance of this factor at different stages of practice on the criterion task, we may again refer to Table 3. It can be seen that the first stage of practice shows an insignificant loading on this factor, but that subsequent stages show a progressive increase to a point later in training, where the loadings become stabilized.

Factor IV appears to be a *Spatial Relations* factor. Most direct evidence for this interpretation is found in the significant loadings of the Discrimination Reaction Time, Instrument Comprehension, Complex Coordination, and Dial and Table Reading Tests, which have consistently been found saturated with this factor. The loadings of the remaining tests on this factor are not inconsistent with the interpretation of this factor as primarily Spatial in nature. A possible exception is the relatively high loading of the Decoding Test. It is, thus, possible that Factor IV may be a composite of Reasoning and Spatial Relations factors. However, the remaining tests are primarily spatial in nature. This factor seems to involve the ability to relate different responses to different stimuli, where either stimuli or responses are arranged in spatial order. Emphasis in this Spatial factor appears to be on decision as to direction of movement.

It can be seen in Table 3 that this factor is most importantly involved in the first stage of practice on the Complex Coordination Test and is not found in later stages of practice.

Factor V is defined primarily by the Visual Pursuit, Speed of Identification, and Spatial Orientation Tests, and also by Instrument Comprehension and Dial and Table Reading. This factor is readily defined as *Perceptual Speed.* It involves the rapid comparison of visual forms and the notation of similarities and differences in form and detail.

It is found in the Complex Coordination task only to a low degree with a loading of .25 or above only in certain of the earlier stages of practice.

Factor VI is defined principally by Pattern Comprehension and Mechanical Principles and also by Decoding, Spatial Orientation, and Dial and Table Reading. The tests with the two highest loadings have consistently

TABLE 2

Centroid Factor Loadings of Test and Practice Task Variables*

Test	I	II	III	IV	V	VI	VII	VIII	IX	X	h^2
1. Practice Stage 1, Complex Coordination	79	04	20	-17	-09	-11	-13	13	07	10	76
2. Practice Stage 2, Complex Coordination	83	34	12	03	-07	-05	-16	08	14	-13	90
3. Practice Stage 3, Complex Coordination	85	33	07	01	-08	-07	-05	13	04	-05	87
4. Practice Stage 4, Complex Coordination	83	43	13	13	02	04	-06	-04	-09	-12	93
5. Practice Stage 5, Complex Coordination	80	48	13	10	03	10	-05	-10	-01	01	92
6. Practice Stage 6, Complex Coordination	75	52	07	14	09	13	-02	-09	05	-02	89
7. Practice Stage 7, Complex Coordination	77	48	09	15	06	08	05	05	-06	14	90
8. Practice Stage 8, Complex Coordination	76	48	14	15	07	11	10	04	-02	02	87
9. Numerical Operations - II	49	-35	-21	21	21	19	-11	19	09	-10	59
10. Dial and Table Reading	71	-44	-08	07	22	10	-10	12	14	04	80
11. Mechanical Principles	56	-29	33	-14	23	-04	-12	-10	07	05	62
12. General Mechanics	45	-20	13	-26	36	-15	-06	07	-22	-05	53
13. Speed of Identification	68	-37	02	15	-17	-08	-03	-10	-05	-03	66
14. Pattern Comprehension	65	-35	07	-17	-14	05	-10	-17	11	15	67
15. Visual Pursuit	55	-11	12	10	-18	-08	14	-04	14	-13	43
16. Decoding	62	-41	06	-03	-17	27	-05	17	-12	14	72
17. Instrument Comprehension	57	-25	-08	-17	-09	19	14	15	-04	-15	53
18. Spatial Orientation	64	-31	14	07	-11	-08	20	08	-10	15	63
19. Speed of Marking	60	-23	-14	24	-13	-14	-08	-07	-05	19	57
20. Log Book Accuracy	53	-21	-25	30	05	-22	-08	-18	-04	11	58
21. Rotary Pursuit	60	17	-10	-13	04	-17	15	13	10	08	50
22. Plane Control	47	16	-07	-29	-03	-07	-09	-03	-13	-09	38
23. Discrimination Reaction Time	64	-22	-15	-12	-18	12	-16	03	-07	-10	58
24. Nut and Bolt	41	-07	10	-18	08	-10	15	-09	-03	-11	27
25. Reaction Time	27	24	-39	03	-05	03	10	-11	-19	-04	34
26. Rate of Movement	39	13	-31	-06	-08	06	08	-21	14	13	36
$\Sigma a^2/k$	41	10	03	02	02	02	01	01	01	01	

* Decimal points omitted.

TABLE 3

Rotated Factor Loadings of Test and Practice Task Variables*

Test	I CC	II PC	III RM	IV SR	Factors V P	VI Vz	VII ME	VIII N	IX PS	X Res	h^2
1. Practice Stage 1, Complex Coordination	24	48	10	39	20	38	28	03	22	02	76
2. Practice Stage 2, Complex Coordination	45	62	26	24	26	17	21	04	15	-15	90
3. Practice Stage 3, Complex Coordination	41	60	33	29	27	16	19	03	18	01	88
4. Practice Stage 4, Complex Coordination	62	45	38	20	25	06	25	00	18	-03	93
5. Practice Stage 5, Complex Coordination	65	46	40	11	16	13	22	00	15	01	92
6. Practice Stage 6, Complex Coordination	63	47	43	02	15	12	20	06	09	-01	89
7. Practice Stage 7, Complex Coordination	60	48	40	12	13	13	16	06	17	21	90
8. Practice Stage 8, Complex Coordination	62	47	37	10	22	10	18	04	08	17	88
9. Numerical Operations - II	07	07	05	21	25	03	16	66	09	-05	59
10. Dial and Table Reading	07	13	03	26	32	30	30	60	21	00	79
11. Mechanical Principles	17	06	-14	15	24	41	49	14	20	-02	60
12. General Mechanics	-03	05	00	26	14	06	62	08	16	13	52
13. Speed of Identificatidn	12	01	10	35	47	29	08	21	37	-09	65
14. Pattern Comprehension	07	04	10	33	23	60	19	18	21	-11	60
15. Visual Pursuit	15	20	09	16	50	24	02	04	11	-06	42
16. Decoding	17	-03	01	59	22	36	06	35	10	11	71
17. Instrument Comprehension	02	09	20	46	35	20	16	24	-12	03	52
18. Spatial Orientation	10	07	04	35	45	32	09	12	27	26	61
19. Speed of Marking	10	09	20	23	26	24	-03	26	50	00	56
20. Log Book Accuracy	03	05	27	04	27	12	07	31	55	-05	58
21. Rotary Pursuit	03	48	34	12	21	18	18	05	10	17	51
22. Plane Control	07	25	33	29	04	09	29	-10	07	-09	38
23. Discrimination Reaction Time	08	12	25	52	22	23	13	24	12	-18	58
24. Nut and Bolt	03	08	14	12	29	17	32	-07	03	03	26
25. Reaction Time	06	08	54	08	03	-11	-03	03	07	01	33
26. Rate of Movement	03	17	48	-01	04	28	-03	12	07	-05	36
$\Sigma a^2/k$	10	10	08	08	07	06	06	05	05	01	

* Decimal points omitted.

identified a factor called *Visualization*. It appears to represent the ability to make "mental" manipulations of visual images or to comprehend imaginary movements. It is found, for example, in tests requiring the examinee to imagine the rotation of depicted objects in three-dimensional space, the folding or unfolding of flat patterns, the relative changes in position of objects in space, and the motion of machinery. In the present analysis, the Complex Coordination Test was found to sample this factor only in the first stage of practice.

Factor VII is readily identified from its high loadings in the General Mechanics and Mechanical Principles Tests. It has been called *Mechanical Experience* in previous analyses. In the present study this factor also contributes somewhat to the Nut and Bolt, Plane Control, and Dial and Table Reading Tests.

With respect to the Complex Coordination Test, the factor appears more involved in the early stages of practice, but only to a low degree. Loadings are not significant in the later stages of practice.

Factor VIII is identified from the high loadings of the Numerical Operations and Dial and Table Reading Tests as the *Numerical Facility* factor. The presence of this factor in tests involving the location and use of numbers as well as in those tests involving computation confirms a previous suggestion (6) that the definition of this factor must be broadened to include more than computational facility. This factor was not common to the Complex Coordination Test at any of the stages of practice.

Factor IX is most strongly sampled by the Log Book Accuracy and Speed of Marking Tests. These two tests have previously defined a factor labeled *Psychomotor Speed* (7). It appears, however, that this factor is confined to printed tests which require simple, rapid movements (marking an answer sheet). It is not common to any of the apparatus tests and is distinct from the Rate of Movement factor. The name Psychomotor Speed, thus, appears too broad since many kinds of tasks involve speed. Whether this factor is broader than speed of marking or whether it represents speed in very simple restricted movements as contrasted with the more gross movement involved in the Rate of Movement factor must await additional data.

Factor X is a residual factor.

Factor Pattern in Relation to Practice

The results indicate clearly that considerable changes occurred in the factor structure of the criterion task as practice on the task was continued. It also is evident that these changes were quite systematic. It appears useful to examine the nature of these changes more closely.

The results indicate that the Complex Coordination Test was most complex (showed higher loadings on more factors) during the first stage of practice, and that the test became less complex factorially as practice was

continued. Thus, the last four stages of practice show significant loadings on only three factors, while during the first two stages significant loadings were obtained on seven different factors.

Moreover, the somewhat lower communality (h^2) of Stage 1 (.76) relative to those of the remaining seven stages (.88-.95) suggests the presence of more unexplained variance ($1 - h^2$) at Stage 1 than at the later stages. Since the reliabilities (Hoyt) of each of the eight stages are high and quite constant (.95-.98), the differences between communalities cannot be explained on the basis of increased error variance during the first practice stage. Thus, the presence of still additional factors during this first stage of practice is suspected.

There was also a shift in the *nature* of the factors contributing variance at early and later stages of practice. For example, the factors primarily involved early in practice were Psychomotor Coordination, Spatial Relations, Visualization, and to some extent Mechanical Experience. Perceptual Speed also appeared to a low degree in certain of the early stages. From the fifth through the last stage of practice, the only factors with significant loadings in the test were the Psychomotor Coordination and Rate of Movement factors and a factor common only to the criterion task itself. This suggests that at early stages of practice on the task, certain "non-motor" factors play a major role, but as practice is continued these become less and less important as contributors to individual differences in performance. At the same time, certain "motor" factors increase in importance as practice is continued until these are the only factors that enter to a significant degree at late stages of practice.

Figure 1 presents this shift graphically. The shaded areas represent the percentage of the total variance in performance accounted for by each factor at each of the eight stages of practice. By combining percentages of variance accounted for by various factors, it can be seen quite clearly that during the first stage of practice on the Complex Coordination Test, so-called "non-motor" factors are contributing 46.1 per cent of the variance, while the three "motor" factors contribute but 29.5 per cent. However, as practice continues there is a systematic decrease in "non-motor" variance accompanied by an increase in "motor" variance until approximately stage 5, beyond which little appreciable change occurs. By the final stage of practice, the non-motor factors account for only 10.5 per cent of the variance while the motor factors now account for 74.5 per cent.

Discussion

In general, these changes in factor loadings at different stages of performance on the criterion task indicate that the quantitative pattern of abilities determining differences in goodness of performance changes with practice. In other words, individual differences in performance on the task

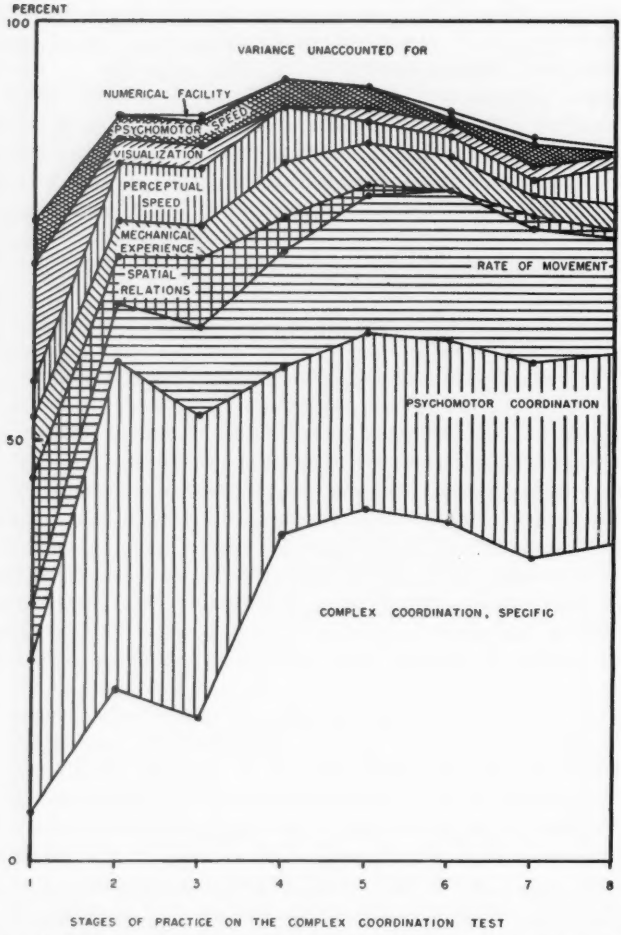


FIG. 1.—Percentage of variance (shaded area) represented by each factor at different stages of practice on the Complex Coordination Test.

after certain amounts of practice are likely to depend more on certain abilities and less on others than they did initially.

Learning theorists have given relatively little attention to the relationships of individual differences to the principles of learning. It is therefore difficult to relate the results of the present study to any specific learning theory. However, a few investigators have given some attention to the problem. Reynolds (10), for example, found decreasing correlations between

printed tests and a psychomotor test as practice was continued on the latter test. He hypothesized that the subject's early performance is to a great extent a function of previous experience and therefore reflects a number of abilities. As learning continues, the subject's performance depends more upon his experience with the particular task employed. To an increasing extent, he learns the specific responses required by the task. The abilities or aptitudes associated with these specific habits being acquired, it is hypothesized, are progressively restricted in extent and generality. This view appears related to Seashore's "work methods" hypothesis (**11**) as well as to some of the findings of Woodrow (cf. **14**). In factor analyses of extended practice on printed tests, Woodrow found that loadings of practice gains scores are often specific to the particular test.

Some of the data of the present study are not inconsistent with these previous interpretations. For example, it has been noted that as practice was continued on the Complex Coordination Test there was a progressive increase in loadings on a factor common only to the test itself. Also consistent with these hypotheses is the shrinkage in the number of factors contributing to performance as practice is continued. However, such "specificity hypotheses" imply that a sizable portion of the between-subjects variance observed at advanced levels of proficiency in a task is not ascribable to variation along any other dimension other than that established by the task itself.

At present, however, this view is regarded as overly pessimistic, especially with respect to the problem of predicting more advanced and terminal proficiency in psychomotor skills. It is felt, rather, that the problem is one of identifying what variables predict such advanced proficiency levels. The present study indicates that it is possible to isolate variables which predict performance at advanced stages. For example, the Rate of Movement and Psychomotor Coordination factors were found to contribute a sizable portion of the variance at advanced stages. The factor found specific to only the trials of the practice task may be broken down still further, with additional experimentation including other variables. Perhaps the most encouraging results regarding the predictability of advanced stages of psychomotor performance are found in recent work by Adams (**1**). He found that a combination of certain test measures yielded a multiple correlation with advanced levels of performance on this criterion task that exceeded the correlation between the first and last stages of practice on the task itself. This gives further indication that the abilities sampled at late stages of practice on a psychomotor task are not necessarily specific to the task, but may represent common factors definable by other test variables.

For the present, performance at any stage of practice is regarded as determined by a *set* of cooperating but independently variable abilities. This view is similar to that proposed by Woodrow (**13**). Continuation of practice on the task does not result in an equal increase or decrease in

favorableness for the operation of all the cooperating determining factors contributing to individual differences. For example, in the early stages of performing the task, learning the spatial relationships of the different stimuli to the different responses may be an important contributor to individual differences. Once this is learned to a sufficiently high degree and no further improvement in this skill is possible (or necessary), within the limits of the task, other features of the task may assume increasing importance (e.g., completing the proper sequence of movements quickly). However, the important point is that earlier in the learning period the use of certain skills is a minimal but necessary requirement for achieving a certain amount of progress in performing the task. At this stage, individual differences in these skills affect the rise in the acquisition curve. Later in training, individual differences in other combinations of skills may be the chief contributors to performance. The rise in the performance curve may be considered to be a resultant of systematic transformations in the particular combination of abilities contributing variance at different stages of practice.

Implications

The results of the present study would appear to have certain implications for test development in this aptitude area and for certain problems of criterion analysis. For example, the fact of changes in factor pattern with practice points up the importance of establishing for such tests what abilities are contributing variance at different stages of performance on the test. The problem appears especially crucial in regard to psychomotor tests which are included as parts of larger, rather comprehensive classification batteries. The need for excluding from such apparatus tests variance measurable by printed tests has been pointed out earlier (**4**). Knowledge about factorial content at different stages should enable one to specify how much practice to give on such tests in the operational test situation in order to reach the stage of performance at which it would be most desirable to score the test; that is, performance up to a given stage would be purely "practice" for the subject and scoring would begin *after* this stage is reached and for a specific period. This would include the stage at which (1) the factors measured by the test which are also measured by the other tests are at a minimum, and (2) the loadings of the remaining valid factors in the test are at a maximum. This would presumably maximize the unique contribution of such tests to the predictive power of the total battery.

The findings of the present study are indirectly related to certain problems of criterion development in operational situations. For example, testing programs are most often evaluated against more immediate criteria of proficiency in lieu of more intermediate or ultimate criteria of performance. Thus, the tests used in the Aircrew Classification Battery were, without exception, designed to predict the success of candidates *in training* for the

specialty to which they were assigned. Similarly, in the industrial situation, test validation is often conducted against proficiency at relatively early stages of training in place of validation against the more elusive on-the-job criteria. Aside from the several reasons supporting the use of the more immediate criteria (and there are many), the data of the present study at least suggest the need for criterion analyses of more advanced levels of proficiency, especially in jobs emphasizing perceptual-motor skills. It is quite conceivable that the abilities contributing to individual differences in earlier stages of skill attainment in such jobs may be somewhat different than those contributing variance at more advanced and terminal levels of proficiency. Thus, the relative size of test regression weights determined against immediate criteria may be expected to shift, when the same battery is evaluated against more advanced criteria. Similarly, tests rejected from the battery because of insignificant validity against immediate criteria may have turned out to possess considerable validity for advanced levels of proficiency. As a corollary of these possibilities, there is a need for developing tests which predict individual differences at more advanced and terminal levels of proficiency in such situations.

REFERENCES

1. Adams, J. A. The prediction of performance at advanced stages of training on a complex psychomotor task. San Antonio, Tex.: Human Resources Research Center, Lackland Air Force Base, December 1953. (*Research Bulletin* 53-49.)
2. Anastasi, A. The influence of specific experience upon mental organization. *Genetic Psychol. Monogr.*, 1936, **18**, 245-255.
3. Fleishman, E. A. A factor analysis of intra-task performance on two psychomotor tests. *Psychometrika*, 1953, **18**, 45-55.
4. Fleishman, E. A. Testing for psychomotor abilities by means of apparatus tests. *Psychol. Bull.*, 1953, **50**, 241-262.
5. Greene, E. B. An analysis of random and systematic changes with practice. *Psychometrika*, 1943, **8**, 37-52.
6. Guilford, J. P., Fruchter, B., and Zimmerman, W. S. Factor analysis of the Army Air Forces Shepard Field Battery of experimental aptitude tests. *Psychometrika*, 1952, **17**, 45-68.
7. Guilford, J. P. (Ed.) Printed Classification Tests. AAF Aviation Psychology Program Research Report No. 5, 1947.
8. Melton, A. W. (Ed.) Apparatus Tests. AAF Aviation Psychology Program Research Report No. 4, 1947.
9. Michael, W. B. Factor analysis of tests and criteria: a comparative study of two AAF pilot populations. *Psychol. Monogr.*, 1949, **63**, No. 298.
10. Reynolds, B. The effect of learning on the predictability of psychomotor performance. *J. exp. Psychol.*, 1952, **44**, 189-198.
11. Seashore, R. H. Work and motor performance. In Stevens, S. S. (Ed.) Handbook of experimental psychology. New York: Wiley, 1951.
12. Thurstone, L. L. Multiple-factor analysis. Chicago: Univ. Chicago Press, 1947.
13. Woodrow, H. Factors in improvement with practice. *J. Psychol.*, 1939, **7**, 55-70.
14. Woodrow, H. The ability to learn. *Psychol. Rev.*, 1946, **53**, 147-158.

15. Woodrow, H. The application of factor-analysis to problems of practice. *J. gen. Psychol.*, 1939, **21**, 457-460.
16. Woodrow, H. The relation between abilities and improvement with practice. *J. educ. Psychol.*, 1938, **29**, 215-230.
17. Woodrow, H. The relation of verbal ability to improvement with practice in verbal tests. *J. educ. Psychol.*, 1939, **30**, 370-382.
18. Zachert, Virginia, and Friedman, G. Factorial comparisons of two Aircrew Classification Batteries with and without the variable of previous flying experience. San Antonio, Tex.: Human Resources Research Center, Lackland Air Force Base, April 1952. (*Research Bulletin* 52-16.)
19. Zachert, Virginia, and Ivens, F. C. Comparative factor analyses of tests using three different scoring techniques. San Antonio, Tex.: Human Resources Research Center, Lackland Air Force Base, August 1951. (*Research Bulletin* 51-16.)
20. Zimmerman, W. S. A simple graphical method for orthogonal rotation of axes. *Psychometrika*, 1946, **11**, 51-55.

Manuscript received 8/8/53
Revised manuscript received 11/19/53

PSYCHOMETRIKA—VOL. 19, NO. 3
SEPTEMBER, 1954

OSGOOD AND SUCI'S MEASURE OF PATTERN SIMILARITY AND *Q*-TECHNIQUE FACTOR ANALYSIS

JAMES M. SAKODA

UNIVERSITY OF CONNECTICUT

The purpose of this article is to show that a relationship can be established between Osgood and Suci's D based on raw scores and scores obtained from factor analysis data without equalizing the means and variances of each individual's set of scores.

The measure of pattern similarity recently suggested by Osgood and Suci (3) deserves further study because it takes into account difference in level, as well as similarity in pattern, of two individual's scores on a set of variables (2). Osgood and Suci used as their "individuals" words such as *death, hero, fate* and as their "tests" such descriptive categories as *high-low, green-red, weak-strong.* They were interested in studying the similarity in the pattern of scores of the different words. They defined D, their measure of pattern similarity, as

$$\sqrt{\sum (X_i - X_l)^2},$$

i.e., the square root of the sum of the squared differences between two individuals' raw scores on a set of variables. It lacks the mathematical sophistication of Mahalonobis' D (**4, 7**), but it has two advantages over it. It is much simpler to calculate, and it is related to Q-technique factor analysis (**1, 5, 6**), which is familiar to many psychologists.

Osgood and Suci demonstrated the existence of a relationship between D_z, their D calculated from standard scores and V. V, the distance between two points in a factor analytic model, was calculated as

$$V = \sqrt{h_i^2 + h_l^2 - 2h_i h_l \cos \phi_{il}}.$$

It could have also been calculated from a table of factor loadings as follows:

$$V = \sqrt{\sum (a_{im} - a_{lm})^2}.$$

They showed that when the communalities of the variables, h_i^2 and h_l^2 were assumed to be equal to unity, then D_z was equal to $\sqrt{k}V$. In demonstrating this relationship between their measure of pattern similarity and a factor analytic model they were forced to use standard scores which equalized the means and standard deviations of different individuals' profiles of scores. It is the purpose of this paper to show that a relationship can be established between Osgood and Suci's D based on raw scores and the factor analytic model.

This opens up a number of interesting prospects for those concerned with pattern analysis and factor analysis.

First, it is convenient to transform Osgood and Suci's D to the following form:

$$\frac{D^2}{k} = \frac{\sum d^2}{k} = \frac{\sum (X_j - X_l)^2}{k},$$

where k is the number of variables and $X_j - X_l$ represents k differences between raw scores of individuals j and l. Let $X_j = x_j + M_j$ and $X_l = x_l + M_l$, where M_j and M_l are means of the k scores for j and l, respectively, and x_j and x_l are deviations from the respective means. Then

$$\frac{\sum d^2}{k} = \frac{\sum [(x_j - x_l) + (M_j - M_l)]^2}{k}$$

$$= \frac{\sum (x_j - x_l)^2}{k} + \frac{k(M_j - M_l)^2}{k} + \frac{2(M_j - M_l)(\sum x_j - \sum x_l)}{k}.$$

The last cross-product term drops out because it contains $\sum x_j$ and $\sum x_l$, both of which are equal to zero. Hence

$$\frac{\sum d^2}{k} = \frac{\sum x_j^2}{k} + \frac{\sum x_l^2}{k} - \frac{2 \sum x_j x_l}{k} + (M_j - M_l)^2$$

$$= \sigma_j^2 + \sigma_l^2 - 2r_{jl}\sigma_j\sigma_l + (M_j - M_l)^2.$$

Now, to obtain comparable results from a Q-technique factor analysis of the same data, it is necessary to take factor loadings on orthogonal axes, a_{jm} and a_{lm}, and transform them to a new set of loadings, b_{jm} and b_{lm}, where

$$b_{jm} = a_{jm}\frac{\sigma_j}{h_j} \quad \text{and} \quad b_{lm} = a_{lm}\frac{\sigma_l}{h_l};$$

σ_j and σ_l are the standard deviations of the scores of individuals j and l, and h_j and h_l are the square roots of the communalities for these individuals. Dividing each factor loading by h normalizes the loadings so that each vector length is equal to unity. Multiplying each factor loading by σ has the effect of making unequal the standard deviations which were equalized in calculating r_{jl} for the correlation matrix. By summing the cross products of factor loadings we would obtain the correlation coefficient, $\sum z_j z_l / N$, between two individuals. By first multiplying the factor loadings by the standard deviation, we would obtain the covariance between two variables:

$$\sum (a_{jm}\sigma_j)(a_{lm}\sigma_l) = r_{jl}\sigma_j\sigma_l = \frac{\sum x_j x_l}{N}.$$

To distinguish this space in which b_{im} and b_{lm} are plotted from the factor analytic model in which factor loadings are plotted we shall call the former a covariance space.

In addition to these transformations a new orthogonal dimension should be added to the existing set of dimensions. The loading of each person on this dimension, which will be labeled M, consists of the mean of the pattern of scores of each individual. We shall refer to the space including this new dimension as the covariance and mean space. W^2, the square of the distance between individuals j and l in the covariance and mean space, is calculated by means of the formula for the distance between two points:

$$W^2 = \sum (b_{im} - b_{lm})^2 + (M_i - M_l)^2$$

$$= \sum \left(a_{im}\frac{\sigma_i}{h_i} - a_{lm}\frac{\sigma_l}{h_l}\right)^2 + (M_i - M_l)^2$$

$$= \frac{\sigma_i^2}{h_i^2}\sum a_{im}^2 + \frac{\sigma_l^2}{h_l^2}\sum a_{lm}^2 - \frac{2\sigma_i\sigma_l}{h_ih_l}\sum a_{im}a_{lm} + (M_i - M_l)^2.$$

Since

$$\sum a_{im}^2 = h_i^2, \qquad \sum a_{lm}^2 = h_l^2 \qquad \text{and} \qquad \sum a_{im}a_{lm} = r_{il},$$

$$W^2 = \sigma_i^2 + \sigma_l^2 - \frac{2r_{il}\sigma_i\sigma_l}{h_ih_l} + (M_i - M_l)^2.$$

When comparing D^2/k and W^2 there is now only one discrepancy. D^2/k contains the term $2r_{il}\sigma_i\sigma_l$, while W^2 contains the term $2r_{il}\sigma_i\sigma_l/h_ih_l$. These terms could be equalized by assuming, as Osgood and Suci did, that h_i^2 and h_l^2 are equal to unity when performing the factor analysis. This assumption would permit the setting up of factor axes in the unique dimensions of each individual. Then it follows that $W^2 = D^2/k$ and $W = D/\sqrt{k}$. Thus it has been shown that Osgood and Suci's D can be obtained from factor analysis without equalizing the means and variances of each individual's set of scores.

The advantage of using communalities rather than unity in factor analysis is that it is possible to deal with a minimum number of common dimensions. If the communalities of the individuals do not vary excessively, the correspondence between W when communalities are not set equal to unity and $D/\sqrt{k}$ will be very close. For Osgood and Suci's example the correlation between W calculated by the above procedure and D was .985. It is probably preferable to use communalities and ignore the difference between W and $D/\sqrt{k}$ than to assume communalities to be equal to unity.

Several implications can be drawn from the correspondence between Osgood and Suci's measure of pattern similarity, D, and W calculated from factor analysis results:

1. If D is redefined as $\sqrt{\sum d^2/k}$, it will approximate the distance be-

tween two individuals in a factor analytic covariance and mean space. This will also prevent D from being a function of the number of variables which are included in the analysis.

2. Those desirous of using Osgood and Suci's D might well consider the advantages accruing from doing a Q-technique factor analysis first and then calculating W's from it. This would involve more calculation, but would permit interpretation of the data via factor analysis, a well-established procedure.

3. A geometric representation of the degree of pattern similarity between individuals, for example, can be obtained by plotting b_{im}'s and means on orthogonal coordinates. This is of practical significance. In this geometric model the distance between two points represents extent of pattern dissimilarity of two individuals' sets of scores. The length of the distance of a point from the origin (i.e., the vector length) represents the standard deviation of a person's set of scores, if the contribution of the additional M axis is not taken into account. This follows since $\sum b_{im}^2 = \sum (\alpha_{im}\sigma_i/h_i)^2 = \sigma_i^2$. The cosine of the angle between two vectors remains equal to $r_{it}/h_i h_t$, if the M axis is not taken into account.

4. Since the M axis is independent of the other orthogonal axes it is possible to study two individuals' profile difference together with, or separately from, differences in means.

5. Those who have performed Q-technique factor analysis in the past and who were dissatisfied with the fact that means and variances of individuals' patterns of scores were equalized can now allow these to vary by transforming factor loadings to b_{im}'s, adding the M axis and calculating distances (W's) between individuals. By clustering together individuals with the smallest W's and seeking clusters with the largest average W's, it is possible to form distinctive groupings of individuals.

REFERENCES

1. Cattell, Raymond B. Factor analysis. New York: Harper, 1952.
2. Gaier, Eugene L. and Lee, Marilyn C. Pattern analysis: the configurational approach to predictive measurement. *Psychol. Bull.*, 1953, **50**, 140–148.
3. Osgood, Charles E. and Suci, George J. A measure of relation determined by both mean difference and profile information. *Psychol. Buli.*, 1952, **49**, 251–262.
4. Rao, C. Radhakrishna. Advanced statistical methods in biometric research. New York: Wiley, 1952.
5. Stephenson, William. The study of behavior, Q-technique and its methodology. Chicago: Univ. Chicago Press, 1953.
6. Thomson, Sir Godfrey. The factorial analysis of human ability. New York: Houghton Mifflin, 1951.
7. Webster, Harold. A note on profile similarity. *Psychol. Bull.*, 1952, **49**, 538–539.

Manuscript received 10/29/53
Revised manuscript received 11/20/53

PSYCHOMETRIKA—VOL. 19, NO. 3
SEPTEMBER, 1954

A COMPUTING CHART FOR THE POINT BISERIAL CORRELATION COEFFICIENT*

HARVEY F. DINGMAN

UNIVERSITY OF SOUTHERN CALIFORNIA

The point biserial correlation coefficient is the preferred statistic when a continuous variable is to be correlated with a variable that is divided into a real dichotomy. The computing chart described here gives a quick estimate of the correlation when a normally distributed continuous variable is artificially dichotomized at its median. The estimate is not biased by non-median dichotomizing of the discontinuous variable. It was designed particularly for use in item analysis. With appropriate modification of data it can be used for other purposes.

In an article by Michael, Perry, and Guilford (2) it is shown that

$$r_{pb} = \frac{y}{\sqrt{pq}} r_t \ , \tag{1}$$

where

p_u = proportion of individuals above the criterion value on variable X and also above the criterion value on variable Y.

p_l = proportion of individuals above the criterion value on variable X but below the criterion value on variable Y.

		Y +	Y −	
X	+	p_u	p_l	$2P$
	−			$2Q$
		1.00	1.00	

$P = \dfrac{p_u + p_l}{2}$

$Q = 1 - P$

X = variable that is a real dichotomy

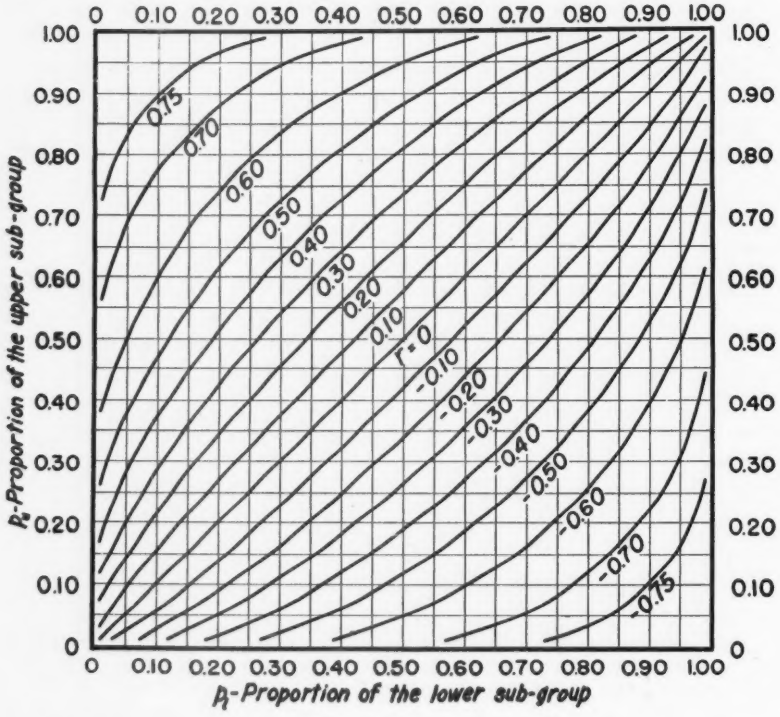
Y = variable that is artificially dichotomized at the median

y = the height of the ordinate of the normal curve at P.

*The writer wishes to acknowledge the help and guidance of J. P. Guilford, who pointed out the need for such a chart and provided financial assistance for the drafting of the chart.

Therefore, to construct a computing chart for the r_{pb} one needs only to construct a computing chart for the tetrachoric correlation and then alter it systematically according to formula (1). Values were selected from the computing diagrams of Chesire, Saffir, and Thurstone (1). A new chart for the tetrachoric r was computed wherein the values tabled were p_u and p_l , when the variable Y is dichotomized at the median.

Along the diagonals of this chart P is a constant as substitution in the above formula will demonstrate. Therefore, $y/\sqrt{pq}$ is also a constant. This convenient fact was utilized to compute the curves on the chart for the point biserial r. Graphical curvilinear interpolation between the lines of the chart for r_t enabled one to find points for the point biserial chart. Then the curves of equal correlation were drawn using french curves.



This chart has been used by Engelhart, who reports in a personal communication as follows:

> For a two-hundred exercise achievement test and for a sample of 200 students we had previously obtained point biserial coefficients of correlation

by means of the second formula given on page 426 of Gulliksen's *Theory of Mental Tests*. We also had the per cent of correct response for each exercise for the hundred students whose total scores were above the median and for the hundred students whose total scores were below the median. These data were used with the new chart by two statistical clerks experienced in using a very similar chart for tetrachoric coefficients of correlation. The estimates were made independently and neither clerk had seen the calculated coefficients for some months. The point biserial r's estimated by one of the clerks correlated .86 with the calculated r's. The other clerk's estimates correlated .88. The two series of estimates correlated .98 (the "reader reliability"). The means of the estimates were the same when rounded off to two decimal places, but were .02 lower than the mean of the r's calculated by the formula. Similarly, the standard deviations of the estimates agreed in the second decimal place, but the standard deviation of the calculated r's was .01 smaller.

This is what should be expected, however. Linear interpolation will tend to produce values that are systematically wrong. The error will increase as the distance increases from the two curves used in interpolation. The correlation between two users of this chart would be high if they use the same method of interpolation. That is, the error in interpolation would be the same for each person.

When the estimates of correlation derived from this chart are compared with correlations derived from raw data, the agreement will be less satisfactory. In the first place, the error in interpolation now will operate to reduce the agreement between the two methods. In the second place, the assumption of normality may not strictly be met (**2**). In some hypothetical data supplied by Engelhart the calculated estimate of correlation was .07 too low because the data were skewed significantly. This tended to increase the sample standard deviation, which reduced the calculated point biserial coefficient. The sample standard deviation is not involved in the use of this chart.

Thus in the use of this chart, as in the use of any formula, the assumptions must be met and the operations done properly if perfectly correct results are to be obtained. However, for results accurate to one decimal place, this chart should be satisfactory under a wide variety of conditions. Larger scale copies of this chart may be obtained from the author free of charge.

REFERENCES

1. Chesire, L., Saffir, M., & Thurstone, L. L. Computing diagrams for the tetrachoric correlation coefficient. Chicago: Univ. Chicago Bookstore, 1933.
2. Michael, W. B., Perry, N. C., & Guilford, J. P. The estimation of a point biserial coefficient of correlation from a phi coefficient. *Brit. J. Psychol., Statis. Sec.*, 1952, **5**, 139-150.

PSYCHOMETRIKA—VOL. 19, NO. 3
SEPTEMBER, 1954

A NOTE ON ITEM ANALYSIS WITH AN ELECTRONIC COMPUTER

JAMES ZIEGLER

TEACHER CHARACTERISTICS STUDY*

An electronic computer with large capacity memory, and employing binary numbers, is ideally suited for large-scale item analysis. The adaptation and use of one computer, SWAC, is described.

The reliability of item analysis data is, in part, a function of the size of the sample upon which the item analysis is based. But the mechanical job in counting reponses to test items, even for a small sample, is time consuming, and for large samples or replicated analyses it often is prohibitively laborious.

The requirements for rapid large-scale item analysis seem to be two: (1) provision for the rapid counting of responses; and (2) availability of a large memory device, easily accessible, in which to store values which are used over and over in computing significances of differences or correlation coefficients.

These requirements could not be met by mechanical equipment previously available, but with the introduction of high speed electronic computers, such as SWAC (Standard's Western Automatic Computer)† the principal obstacles were removed.

The computer SWAC uses the binary system rather than the decimal system. A large capacity high speed electro-magnetic memory is attached. This combination of binary number system and large size high speed memory is readily adaptable to satisfy the two major requirements previously referred to for large-scale item analysis.

In using SWAC for item analysis, an adaptation of Flanagan's Table of Biserial Correlation Coefficients (1), which provides estimates of product-moment r's based upon the responses of upper and lower 27% criterion groups, was developed. The column and row headings, representing fractions of the upper and lower criterion groups, were divided into 127ths (giving

*A research project of the American Council on Education and The Grant Foundation.

†The Institute for Numerical Analysis, a branch of the National Bureau of Standards, located on the campus of the University of California at Los Angeles, has designed and constructed this computer which, by arrangement with the University, has been made available to research groups located on the campus. The Teacher Characteristics Study has been one of the research groups privileged to employ the facilities of SWAC.

128 entries in any one row or column), rather than fractions of 100. The adaptation was made assuming linearity between two adjacent entries in the original Flanagan table. This arrangement simplified operation of the computer and also increased computation speed.

With this preliminary step completed, item analysis could be undertaken with SWAC performing the necessary operations as directed. IBM cards, with item responses punched into them, were fed into SWAC and summary cards periodically punched out. It was possible to count a maximum of 34 five-response items at a time, the size of the sample being restricted to 510 in each of the high and low criterion groups and 1020 in the middle group. If any group exceeded these limitations, separate counts were made for each group by introducing a slight modification in the basic coding.

Summary cards were automatically punched out by the computer, showing for each item response the frequency of selection of the response by members of the high, middle, and low criterion groups, and of the sample as a whole. One of the summary cards also indicated the number of members in each of the three criterion groups and the total N.

Next, the adaptation of Flanagan's correlation table was stored in the electro-magnetic memory, and the summary cards from the previous operation were fed into SWAC. Final results were automatically punched out on IBM cards and then listed by IBM Tabulator to provide a printed record of the several values. The following frequently used item analysis data were printed for each item response:

(1) per cent of each criterion group (high, middle, and low) selecting the response;
(2) per cent of the entire sample (the three groups combined) selecting the response; and
(3) estimated biserial correlation of the response with the criterion.

Among the advantages to be gained in the use of such equipment are (1) the possibility of enlargement of the original pool of items to many times the present limit, and (2) the possibility of employing substantially larger samples of respondents in item validity research. Even with increased pools of items and much larger subject samples, the labor involved in item analysis consumes but a fraction of the time required by other methods.

REFERENCE

1. Flanagan, J. C. General considerations in the selection of test items and a short method of estimating the product-moment coefficients from data at the tails of the distribution. *J. educ. Psych.* 1939, **30**, 674-680.

Manuscript received 9/11/53

Revised manuscript received 10/28/53